

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (OTHER THAN US): FLORIGENE LIMITED

(US ONLY): Filippa BRUGLIERA, Timothy Albert HOLTON, Michael Zeron MICHAEL

(ii) TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID PATHWAY ENZYMES AND USES THEREFOR

(iii) NUMBER OF SEQUENCES: 40

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE

(B) STREET: 1 LITTLE COLLINS STREET

(C) CITY: MELBOURNE

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

- 110 -

(B) FILING DATE: 28-FEB-1997

(C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PN8386

(B) FILING DATE: 28-FEB-1997

(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES, DR E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/AF

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

(C) TELEX: AA 31787

- 111 -

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1789 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 50..1586

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAGGAATTG GTGAACCCCA TAGAAGTAAA ATACTCCTAT CTTTATTTC ATG GAA	55
Met Glu	
1	
ATC TTA AGC CTA ATT CTG TAC ACC GTC ATT TTC TCA TTT CTT CTA CAA	103
Ile Leu Ser Leu Ile Leu Tyr Thr Val Ile Phe Ser Phe Leu Leu Gln	
5 10 15	
TTC ATT CTT AGA TCA TTT TTC CGT AAA CGT TAC CCT TTA CCA TTA CCA	151
Phe Ile Leu Arg Ser Phe Phe Arg Lys Arg Tyr Pro Leu Pro Leu Pro	
20 25 30	
CCA GGT CCA AAA CCA TGG CCA ATT ATA GGA AAC CTA GTC CAT CTT GGA	199
Pro Gly Pro Lys Pro Trp Pro Ile Ile Gly Asn Leu Val His Leu Gly	
35 40 45 50	
CCC AAA CCA CAT CAA TCA ACT GCA GCC ATG GCT CAA ACT TAT GGA CCA	247
Pro Lys Pro His Gln Ser Thr Ala Ala Met Ala Gln Thr Tyr Gly Pro	
55 60 65	

- 112 -

CTC ATG TAT CTT AAG ATG GGG TTC GTA GAC GTG GTG GTT GCA GCC TCG	295
Leu Met Tyr Leu Lys Met Gly Phe Val Asp Val Val Val Ala Ala Ser	
70 75 80	
GCA TCG GTT GCA GCT CAG TTC TTG AAA ACT CAT GAT GCT AAT TTC TCG	343
Ala Ser Val Ala Ala Gln Phe Leu Lys Thr His Asp Ala Asn Phe Ser	
85 90 95	
AGC CGT CCA CCA AAT TCT GGT GCA GAA CAT ATG GCT TAT AAT TAT CAG	391
Ser Arg Pro Pro Asn Ser Gly Ala Glu His Met Ala Tyr Asn Tyr Gln	
100 105 110	
GAT CTT GTT TTT GCA CCT TAT GGA CCT AGA TGG CGT ATG CTT AGG AAA	439
Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu Arg Lys	
115 120 125 130	
ATT TGC TCA GTT CAC CTT TTC TCT ACC AAG GCT TTA GAT GAC TTC CGC	487
Ile Cys Ser Val His Leu Phe Ser Thr Lys Ala Leu Asp Asp Phe Arg	
135 140 145	
CAT GTC CGC CAG GAT GAA GTG AAA ACA CTG ACG CGC GCA CTA GCA AGT	535
His Val Arg Gln Asp Glu Val Lys Thr Leu Thr Arg Ala Leu Ala Ser	
150 155 160	
GCA GGC CAA AAG CCA GTC AAA TTA GGT CAG TTA TTG AAC GTG TGC ACG	583
Ala Gly Gln Lys Pro Val Lys Leu Gly Gln Leu Leu Asn Val Cys Thr	
165 170 175	
ACG AAC GCA CTC GCG CGA GTA ATG CTA GGT AAG CGA GTA TTT GCC GAC	631
Thr Asn Ala Leu Ala Arg Val Met Leu Gly Lys Arg Val Phe Ala Asp	
180 185 190	
GGA AGT GGC GAT GTT GAT CCA CAA GCG GCG GAG TTC AAG TCA ATG GTG	679
Gly Ser Gly Asp Val Asp Pro Gln Ala Ala Glu Phe Lys Ser Met Val	
195 200 205 210	
GTG GAA ATG ATG GTA GTC GCC GGT GTT TTT AAC ATT GGT GAT TTT ATT	727

112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

- 113 -

Val Glu Met Met Val Val Ala Gly Val Phe Asn Ile Gly Asp Phe Ile	
215 220 225	
CCG CAA CTT AAT TGG TTA GAT ATT CAA GGT GTA GCC GCT AAA ATG AAG	775
Pro Gln Leu Asn Trp Leu Asp Ile Gln Gly Val Ala Ala Lys Met Lys	
230 235 240	
AAG CTC CAC GCG CGT TTC GAC GCG TTC TTG ACT GAT ATA CTT GAA GAG	823
Lys Leu His Ala Arg Phe Asp Ala Phe Leu Thr Asp Ile Leu Glu Glu	
245 250 255	
CAT AAG GGT AAA ATT TTT GGA GAA ATG AAA GAT TTG TTG AGT ACT TTG	871
His Lys Gly Lys Ile Phe Gly Glu Met Lys Asp Leu Leu Ser Thr Leu	
260 265 270	
ATC TCT CTT AAA AAT GAT GAT GCG GAT AAT GAT GGA GGG AAA CTC ACT	919
Ile Ser Leu Lys Asn Asp Asp Ala Asp Asn Asp Gly Gly Lys Leu Thr	
275 280 285 290	
GAT ACA GAA ATT AAA GCA TTA CTT TTG AAC TTG TTT GTA GCT GGA ACA	967
Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Val Ala Gly Thr	
295 300 305	
GAC ACA TCT TCT AGT ACA GTT GAA TGG GCC ATT GCT GAG CTT ATT CGT	1015
Asp Thr Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg	
310 315 320	
AAT CCA AAA ATA CTA GCC CAA GCC CAG CAA GAG ATC GAC AAA GTC GTT	1063
Asn Pro Lys Ile Leu Ala Gln Ala Gln Gln Glu Ile Asp Lys Val Val	
325 330 335	
GGA AGG GAC CGG CTA GTT GGC GAA TTG GAC CTA GCC CAA TTG ACA TAC	1111
Gly Arg Asp Arg Leu Val Gly Glu Leu Asp Leu Ala Gln Leu Thr Tyr	
340 345 350	
TTG GAA GCT ATA GTC AAG GAA ACC TTT CGG CTT CAT CCA TCA ACC CCT	1159
Leu Glu Ala Ile Val Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro	

- 115 -

GAGCAGCAAC AGCCCATGGA GATAACATGA GTGTAAATG TATGAGTCTC CATATCTTGT 1646

TTAGTTTGTG TATGCTTTGG ATTTAGTAGT TTTTATATTG ATAGATCAAT GTTGCATTG 1706

TCAGTAAGAA TATCCGTTGC TTGTTTCATT AACTCCAGGT GGACAATAAA AGAAGTAATT 1766

TGTATGAAAA AAAAAAAAAA AAA 1789

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 512 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Glu	Ile	Leu	Ser	Leu	Ile	Leu	Tyr	Thr	Val	Ile	Phe	Ser	Phe	Leu	1	5	10	15
Leu	Gln	Phe	Ile	Leu	Arg	Ser	Phe	Phe	Arg	Lys	Arg	Tyr	Pro	Leu	Pro	20	25	30	
Leu	Pro	Pro	Gly	Pro	Lys	Pro	Trp	Pro	Ile	Ile	Gly	Asn	Leu	Val	His	35	40	45	
Leu	Gly	Pro	Lys	Pro	His	Gln	Ser	Thr	Ala	Ala	Met	Ala	Gln	Thr	Tyr	50	55	60	
Gly	Pro	Leu	Met	Tyr	Leu	Lys	Met	Gly	Phe	Val	Asp	Val	Val	Val	Ala	65	70	75	80
Ala	Ser	Ala	Ser	Val	Ala	Ala	Gln	Phe	Leu	Lys	Thr	His	Asp	Ala	Asn	85	90	95	

- 116 -

Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Glu His Met Ala Tyr Asn
 100 105 110

Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu
 115 120 125

Arg Lys Ile Cys Ser Val His Leu Phe Ser Thr Lys Ala Leu Asp Asp
 130 135 140

Phe Arg His Val Arg Gln Asp Glu Val Lys Thr Leu Thr Arg Ala Leu
 145 150 155 160

Ala Ser Ala Gly Gln Lys Pro Val Lys Leu Gly Gln Leu Leu Asn Val
 165 170 175

Cys Thr Thr Asn Ala Leu Ala Arg Val Met Leu Gly Lys Arg Val Phe
 180 185 190

Ala Asp Gly Ser Gly Asp Val Asp Pro Gln Ala Ala Glu Phe Lys Ser
 195 200 205

Met Val Val Glu Met Met Val Val Ala Gly Val Phe Asn Ile Gly Asp
 210 215 220

Phe Ile Pro Gln Leu Asn Trp Leu Asp Ile Gln Gly Val Ala Ala Lys
 225 230 235 240

Met Lys Lys Leu His Ala Arg Phe Asp Ala Phe Leu Thr Asp Ile Leu
 245 250 255

Glu Glu His Lys Gly Lys Ile Phe Gly Glu Met Lys Asp Leu Leu Ser
 260 265 270

Thr Leu Ile Ser Leu Lys Asn Asp Asp Ala Asp Asn Asp Gly Gly Lys
 275 280 285

Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Val Ala

116

- 117 -

290	295	300
Gly Thr Asp Thr Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu		
305	310	315 320
Ile Arg Asn Pro Lys Ile Leu Ala Gln Ala Gln Gln Glu Ile Asp Lys		
325	330	335
Val Val Gly Arg Asp Arg Leu Val Gly Glu Leu Asp Leu Ala Gln Leu		
340	345	350
Thr Tyr Leu Glu Ala Ile Val Lys Glu Thr Phe Arg Leu His Pro Ser		
355	360	365
Thr Pro Leu Ser Leu Pro Arg Ile Ala Ser Glu Ser Cys Glu Ile Asn		
370	375	380
Gly Tyr Phe Ile Pro Lys Gly Ser Thr Leu Leu Leu Asn Val Trp Ala		
385	390	395 400
Ile Ala Arg Asp Pro Asn Ala Trp Ala Asp Pro Leu Glu Phe Arg Pro		
405	410	415
Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Lys Val Asp Val Arg Gly		
420	425	430
Asn Asp Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala		
435	440	445
Gly Met Asn Leu Gly Ile Arg Met Val Gln Leu Met Ile Ala Thr Leu		
450	455	460
Ile His Ala Phe Asn Trp Asp Leu Val Ser Gly Gln Leu Pro Glu Met		
465	470	475 480
Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Asp Pro		
485	490	495

- 118 -

Leu Val Val His Pro Arg Pro Arg Leu Glu Ala Gln Ala Tyr Ile Gly
 500 505 510

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1745 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 172..1660

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AAGTTCGGCA CGAGCGTCAC ATTCACACCG TCACATTACT ATTCAAACCA CTCATTTTCT 60

ACCTCTCTTT TCTACCCACC AAAACAAAAC AAAACAAAAA AAAACACATA AAAAAACTCA 120

AAAAAAAATT ATAATGTCAC CCTTAGAGGT AACTTTCTAC ACCATAGTCC T ATG CAC 177
 Met His
 1

AAT CTC TAC TAC CTC ATC ACC ACC GTC TTC CGC GGC CAC CAA AAA CCG 225
 Asn Leu Tyr Tyr Leu Ile Thr Thr Val Phe Arg Gly His Gln Lys Pro
 5 10 15

CTT CCT CCA GGG CCA CGA CCA TGG CCC ATC GTG GGA AAC CTC CCA CAT 273
 Leu Pro Pro Gly Pro Arg Pro Trp Pro Ile Val Gly Asn Leu Pro His
 20 25 30

- 119 -

ATG GGC CAG GCA CCG CAC CAG GGC TTA GCA GCC CTG GCG CAA AAG TAT	321
Met Gly Gln Ala Pro His Gln Gly Leu Ala Ala Leu Ala Gln Lys Tyr	
35 40 45 50	
GGC CCT CTA TTG TAT ATG AGA CTG GGG TAC GTG GAC GTT GTT GTG GCC	369
Gly Pro Leu Leu Tyr Met Arg Leu Gly Tyr Val Asp Val Val Val Ala	
55 60 65	
GCC TCA GCG TCT GTA GCG ACC CAG TTT CTT AAG ACA CAT GAC CTA AAT	417
Ala Ser Ala Ser Val Ala Thr Gln Phe Leu Lys Thr His Asp Leu Asn	
70 75 80	
TTT TCG AGT AGG CCA CCG AAT TCG GGG GCT AAA CAC ATT GCT TAT AAC	465
Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Lys His Ile Ala Tyr Asn	
85 90 95	
TAT CAA GAC CTT GTT TTT GCA CCT TAT GGA CCT AAA TGG CGC ATG CTT	513
Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Lys Trp Arg Met Leu	
100 105 110	
AGG AAA ATT TGT TCC TTA CAC ATG TTT TCT TCT AAG GCT TTG GAC GAT	561
Arg Lys Ile Cys Ser Leu His Met Phe Ser Ser Lys Ala Leu Asp Asp	
115 120 125 130	
TTT AGA CTT GTC CGT CAG GAA GAA GTA TCT ATA CTG GTA AAT GCG ATA	609
Phe Arg Leu Val Arg Gln Glu Glu Val Ser Ile Leu Val Asn Ala Ile	
135 140 145	
GCA AAA GCA GGA ACA AAG CCA GTA CAA CTA GGA CAA CTA CTC AAC GTG	657
Ala Lys Ala Gly Thr Lys Pro Val Gln Leu Gly Gln Leu Leu Asn Val	
150 155 160	
TGC ACC ACA AAT GCC TTA TCG AGG GTG ATG CTA GGG AAG CGA GTT CTC	705
Cys Thr Thr Asn Ala Leu Ser Arg Val Met Leu Gly Lys Arg Val Leu	
165 170 175	
GGT GAT GGC ACA GGG AAA AGC GAC CCA AAA GCC GAG GAA TTT AAG GAC	753

- 120 -

Gly Asp Gly Thr Gly Lys Ser Asp Pro Lys Ala Glu Glu Phe Lys Asp	
180 185 190	
ATG GTG CTG GAG TTA ATG GTT CTC ACC GGA GTT TTT AAC ATT GGC GAT	801
Met Val Leu Glu Leu Met Val Leu Thr Gly Val Phe Asn Ile Gly Asp	
195 200 205 210	
TTT GTA CCG GCA TTG GAA TGT CTA GAC TTA CAA GGT GTT GCA TCT AAA	849
Phe Val Pro Ala Leu Glu Cys Leu Asp Leu Gln Gly Val Ala Ser Lys	
215 220 225	
ATG AAG AAA TTA CAT AAA AGA CTT GAT AAT TTT ATG AGT AAC ATT TTG	897
Met Lys Lys Leu His Lys Arg Leu Asp Asn Phe Met Ser Asn Ile Leu	
230 235 240	
GAG GAA CAC AAG AGT GTT GCA CAT CAA CAA AAT GGT GGA GAT TTG CTA	945
Glu Glu His Lys Ser Val Ala His Gln Gln Asn Gly Gly Asp Leu Leu	
245 250 255	
AGC ATT TTG ATA TCT TTG AAG GAT AAT TGT GAT GGT GAA GGT GGC AAG	993
Ser Ile Leu Ile Ser Leu Lys Asp Asn Cys Asp Gly Glu Gly Gly Lys	
260 265 270	
TTT AGT GCC ACA GAA ATT AAG GCC TTG CTA TTG GAT TTA TTT ACA GCT	1041
Phe Ser Ala Thr Glu Ile Lys Ala Leu Leu Leu Asp Leu Phe Thr Ala	
275 280 285 290	
GGA ACA GAC ACA TCA TCT AGT ACA ACT GAA TGG GCC ATA GCC GAA CTA	1089
Gly Thr Asp Thr Ser Ser Ser Thr Thr Glu Trp Ala Ile Ala Glu Leu	
295 300 305	
ATT CGC CAT CCA AAA ATC TTA GCC CAA GTT CAA CAA GAA ATG GAC TCA	1137
Ile Arg His Pro Lys Ile Leu Ala Gln Val Gln Gln Glu Met Asp Ser	
310 315 320	
GTC GTG GGC CGA GAC CGA CTC ATA GCC GAA GCT GAC ATA CCG AAC CTA	1185
Val Val Gly Arg Asp Arg Leu Ile Ala Glu Ala Asp Ile Pro Asn Leu	

- 121 -

325	330	335	
ACC TAC TTC CAA GCC GTA ATC AAA GAG GTT TTC CGA CTT CAC CCG TCC			1233
Thr Tyr Phe Gln Ala Val Ile Lys Glu Val Phe Arg Leu His Pro Ser			
340	345	350	
ACC CCG CTT TCC CTA CCA CGG GTC GCA AAC GAA TCG TGC GAA ATA AAC			1281
Thr Pro Leu Ser Leu Pro Arg Val Ala Asn Glu Ser Cys Glu Ile Asn			
355	360	365	370
GGG TAC CAC ATT CCC AAA AAC ACC ACT TTA TTG GTA AAT GTG TGG GCC			1329
Gly Tyr His Ile Pro Lys Asn Thr Thr Leu Leu Val Asn Val Trp Ala			
375	380	385	
ATC GCA CGC GAC CCT GAG GTT TGG GCC GAC CCG TTA GAG TTT AAA CCC			1377
Ile Ala Arg Asp Pro Glu Val Trp Ala Asp Pro Leu Glu Phe Lys Pro			
390	395	400	
GAA AGA TTT TTG CCG GGC GGC GAA AAG CCC AAT GTG GAT GTG AAA GGA			1425
Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val Asp Val Lys Gly			
405	410	415	
AAC GAT TTT GAG CTG ATT CCG TTC GGG GCG GGC CGA CGG ATT TGT GCT			1473
Asn Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala			
420	425	430	
GGG CTG AGT TTG GGC CTG CGT ATG GTC CAG TTG ATG ACA GCC ACT TTG			1521
Gly Leu Ser Leu Gly Leu Arg Met Val Gln Leu Met Thr Ala Thr Leu			
435	440	445	450
GCC CAT ACT TAT GAT TGG GCC TTA GCT GAT GGG CTT ATG CCC GAA AAG			1569
Ala His Thr Tyr Asp Trp Ala Leu Ala Asp Gly Leu Met Pro Glu Lys			
455	460	465	
CTT AAC ATG GAT GAG GCT TAT GGG CTT ACC TTA CAG CGT AAG GTG CCA			1617
Leu Asn Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln Arg Lys Val Pro			
470	475	480	

- 122 -

CTT AAT GGT CCA CCC GAC CCC GTC GGC TTC TCG GCC CGT GTT T 1660
 Leu Asn Gly Pro Pro Asp Pro Val Gly Phe Ser Ala Arg Val
 485 490 495

AATAATTCCG GGGTTTTTAA AAGCGGGTTA CTTTGTGTTA TGTATTATTC CGTACTAGTT 1720

TGAAAATAAT GGTATTAGAG AAATG 1745

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 496 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met His Asn Leu Tyr Tyr Leu Ile Thr Thr Val Phe Arg Gly His Gln
 1 5 10 15
 Lys Pro Leu Pro Pro Gly Pro Arg Pro Trp Pro Ile Val Gly Asn Leu
 20 25 30
 Pro His Met Gly Gln Ala Pro His Gln Gly Leu Ala Ala Leu Ala Gln
 35 40 45
 Lys Tyr Gly Pro Leu Leu Tyr Met Arg Leu Gly Tyr Val Asp Val Val
 50 55 60
 Val Ala Ala Ser Ala Ser Val Ala Thr Gln Phe Leu Lys Thr His Asp
 65 70 75 80
 Leu Asn Phe S r Ser Arg Pro Pro Asn Ser Gly Ala Lys His Ile Ala
 85 90 95

- 123 -

Tyr Asn Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Lys Trp Arg
100 105 110

Met Leu Arg Lys Ile Cys Ser Leu His Met Phe Ser Ser Lys Ala Leu
115 120 125

Asp Asp Phe Arg Leu Val Arg Gln Glu Glu Val Ser Ile Leu Val Asn
130 135 140

Ala Ile Ala Lys Ala Gly Thr Lys Pro Val Gln Leu Gly Gln Leu Leu
145 150 155 160

Asn Val Cys Thr Thr Asn Ala Leu Ser Arg Val Met Leu Gly Lys Arg
165 170 175

Val Leu Gly Asp Gly Thr Gly Lys Ser Asp Pro Lys Ala Glu Glu Phe
180 185 190

Lys Asp Met Val Leu Glu Leu Met Val Leu Thr Gly Val Phe Asn Ile
195 200 205

Gly Asp Phe Val Pro Ala Leu Glu Cys Leu Asp Leu Gln Gly Val Ala
210 215 220

Ser Lys Met Lys Lys Leu His Lys Arg Leu Asp Asn Phe Met Ser Asn
225 230 235 240

Ile Leu Glu Glu His Lys Ser Val Ala His Gln Gln Asn Gly Gly Asp
245 250 255

Leu Leu Ser Ile Leu Ile Ser Leu Lys Asp Asn Cys Asp Gly Glu Gly
260 265 270

Gly Lys Phe Ser Ala Thr Glu Ile Lys Ala Leu Leu Leu Asp Leu Phe
275 280 285

Thr Ala Gly Thr Asp Thr Ser Ser Ser Thr Thr Glu Trp Ala Ile Ala

- 124 -

290	295	300
Glu Leu Ile Arg His Pro Lys Ile Leu Ala Gln Val Gln Gln Glu Met		
305	310	315 320
Asp Ser Val Val Gly Arg Asp Arg Leu Ile Ala Glu Ala Asp Ile Pro		
325	330	335
Asn Leu Thr Tyr Phe Gln Ala Val Ile Lys Glu Val Phe Arg Leu His		
340	345	350
Pro Ser Thr Pro Leu Ser Leu Pro Arg Val Ala Asn Glu Ser Cys Glu		
355	360	365
Ile Asn Gly Tyr His Ile Pro Lys Asn Thr Thr Leu Leu Val Asn Val		
370	375	380
Trp Ala Ile Ala Arg Asp Pro Glu Val Trp Ala Asp Pro Leu Glu Phe		
385	390	395 400
Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val Asp Val		
405	410	415
Lys Gly Asn Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile		
420	425	430
Cys Ala Gly Leu Ser Leu Gly Leu Arg Met Val Gln Leu Met Thr Ala		
435	440	445
Thr Leu Ala His Thr Tyr Asp Trp Ala Leu Ala Asp Gly Leu Met Pro		
450	455	460
Glu Lys Leu Asn Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln Arg Lys		
465	470	475 480
Val Pro Leu Asn Gly Pro Pro Asp Pro Val Gly Phe Ser Ala Arg Val		
485	490	495

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

- 125 -

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1711 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 91..1629

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

CGAATTCCCC CCCCCCACA CCATTCAATG CCTAAGTCCT CCATTGCGG GCCTAATAAC      60

TAAAAGCCCA CTCTTCCGA CCATCTATAC ATG CAA CAC CAA TAT TAT TCT TTA      114
      Met Gln His Gln Tyr Tyr Ser Leu
                1                5

ATT ACG ATG GAT GAT ATT AGC ATA ACC AGC TTA TTG GTG CCA TGT ACT      162
Ile Thr Met Asp Asp Ile Ser Ile Thr Ser Leu Leu Val Pro Cys Thr
      10                15                20

TTT ATA TTA GGG TTC TTG CTT CTA TAT TCC TTC CTC AAC AAA AAA GTA      210
Phe Ile Leu Gly Phe Leu Leu Leu Tyr Ser Phe Leu Asn Lys Lys Val
      25                30                35                40

AAG CCA CTG CCA CCT GGA CCG AAG CCA TGG CCC ATC GTC GGA AAT CTG      258
Lys Pro Leu Pro Pro Gly Pro Lys Pro Trp Pro Ile Val Gly Asn Leu
                45                50                55

CCA CAT CTT GGG CCG AAG CCC CAC CAG TCG ATG GCG GCG CTG GCA CGG      306

```

- 126 -

Pro His Leu Gly Pro Lys Pro His Gln Ser Met Ala Ala Leu Ala Arg	
60 65 70	
GTG CAC GGC CCA TTA ATT CAT CTG AAG ATG GGC TTT GTG CAT GTG GTT	354
Val His Gly Pro Leu Ile His Leu Lys Met Gly Phe Val His Val Val	
75 80 85	
GTG GCC TCC TCA GCA TCC GTT GCG GAG AAA TTT CTG AAG GTG CAT GAC	402
Val Ala Ser Ser Ala Ser Val Ala Glu Lys Phe Leu Lys Val His Asp	
90 95 100	
GCA AAC TTC TCG AGC AGG CCT CCC AAT TCG GGT GCA AAA CAC GTG GCC	450
Ala Asn Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala	
105 110 115 120	
TAC AAC TAT CAG GAC TTG GTC TTT GCT CCT TAT GGC CCA CGC TGG CGG	498
Tyr Asn Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg	
125 130 135	
ATG CTC AGG AAA ATC TGT GCA CTC CAC CTC TTC TCC GCC AAA GCC TTG	546
Met Leu Arg Lys Ile Cys Ala Leu His Leu Phe Ser Ala Lys Ala Leu	
140 145 150	
AAC GAC TTC ACA CAC GTC AGA CAG GAT GAG GTG GGG ATC CTC ACT CGC	594
Asn Asp Phe Thr His Val Arg Gln Asp Glu Val Gly Ile Leu Thr Arg	
155 160 165	
GTT CTA GCA GAT GCA GGA GAA ACG CCG TTG AAA TTA GGG CAG ATG ATG	642
Val Leu Ala Asp Ala Gly Glu Thr Pro Leu Lys Leu Gly Gln Met Met	
170 175 180	
AAC ACA TGC GCC ACC AAT GCA ATA GCG CGT GTT ATG TTG GGT CGA CGC	690
Asn Thr Cys Ala Thr Asn Ala Ile Ala Arg Val Met Leu Gly Arg Arg	
185 190 195 200	
GTG GTT GGA CAC GCA GAC TCA AAG GCG GAG GAG TTT AAG GCA ATG GTA	738
Val Val Gly His Ala Asp Ser Lys Ala Glu Glu Phe Lys Ala Met Val	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

- 127 -

205	210	215	
GTG GAG TTG ATG GTA TTA GCT GGT GTG TTC AAC TTA GGT GAT TTT ATC			786
Val Glu Leu Met Val Leu Ala Gly Val Phe Asn Leu Gly Asp Phe Ile			
220	225	230	
CCA CCT CTT GAA AAA TTG GAT CTT CAA GGT GTC ATT GCT AAG ATG AAG			834
Pro Pro Leu Glu Lys Leu Asp Leu Gln Gly Val Ile Ala Lys Met Lys			
235	240	245	
AAG CTT CAC TTG CGT TTC GAC TCG TTC TTG AGT AAG ATC CTT GGA GAC			882
Lys Leu His Leu Arg Phe Asp Ser Phe Leu Ser Lys Ile Leu Gly Asp			
250	255	260	
CAC AAG ATC AAC AGC TCA GAT GAA ACC AAA GGC CAT TCG GAT TTG TTG			930
His Lys Ile Asn Ser Ser Asp Glu Thr Lys Gly His Ser Asp Leu Leu			
265	270	275	280
AAC ATG TTA ATT TCT TTG AAG GAC GCT GAT GAT GCC GAA GGA GGG AGG			978
Asn Met Leu Ile Ser Leu Lys Asp Ala Asp Asp Ala Glu Gly Gly Arg			
285	290	295	
CTC ACC GAC GTA GAA ATT AAA GCG TTG CTC TTG AAC TTG TTT GCT GCA			1026
Leu Thr Asp Val Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ala Ala			
300	305	310	
GGA ACT GAC ACA ACA TCA AGC ACT GTG GAA TGG TGC ATA GCT GAG TTA			1074
Gly Thr Asp Thr Thr Ser Ser Thr Val Glu Trp Cys Ile Ala Glu Leu			
315	320	325	
GTA CGA CAT CCT GAA ATC CTT GCC CAA GTC CAA AAA GAA CTC GAC TCT			1122
Val Arg His Pro Glu Ile Leu Ala Gln Val Gln Lys Glu Leu Asp Ser			
330	335	340	
GTT GTT GGT AAG AAT CGG GTG GTG AAG GAG GCT GAT CTG GCC GGA TTA			1170
Val Val Gly Lys Asn Arg Val Val Lys Glu Ala Asp Leu Ala Gly Leu			
345	350	355	360

1170
 1122
 1074
 1026
 978
 930
 882
 834
 786

- 128 -

CCA TTC CTC CAA GCG GTC GTC AAG GAA AAT TTC CGA CTC CAT CCC TCC	1218
Pro Phe Leu Gln Ala Val Val Lys Glu Asn Phe Arg Leu His Pro Ser	
365 370 375	
ACC CCG CTC TCC CTA CCG AGG ATC GCA CAT GAG AGT TGT GAA GTG AAT	1266
Thr Pro Leu Ser Leu Pro Arg Ile Ala His Glu Ser Cys Glu Val Asn	
380 385 390	
GGA TAC TTG ATT CCA AAG GGT TCG ACA CTT CTT GTC AAT GTT TGG GCA	1314
Gly Tyr Leu Ile Pro Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala	
395 400 405	
ATT GCT CGC GAT CCA AAT GTG TGG GAT GAA CCA CTA GAG TTC CGG CCT	1362
Ile Ala Arg Asp Pro Asn Val Trp Asp Glu Pro Leu Glu Phe Arg Pro	
410 415 420	
GAA CGA TTC TTG AAG GGC GGG GAA AAG CCT AAT GTC GAT GTT AGA GGG	1410
Glu Arg Phe Leu Lys Gly Gly Glu Lys Pro Asn Val Asp Val Arg Gly	
425 430 435 440	
AAT GAT TTC GAA TTG ATA CCG TTC GGA GCG GGC CGA AGA ATT TGT GCA	1458
Asn Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala	
445 450 455	
GGA ATG AGC TTA GGA ATA CGT ATG GTC CAG TTG TTG ACA GCA ACT TTG	1506
Gly Met Ser Leu Gly Ile Arg Met Val Gln Leu Leu Thr Ala Thr Leu	
460 465 470	
AAC CAT GCG TTT GAC TTT GAT TTG GCG GAT GGA CAG TTG CCT GAA AGC	1554
Asn His Ala Phe Asp Phe Asp Leu Ala Asp Gly Gln Leu Pro Glu Ser	
475 480 485	
TTA AAC ATG GAG GAA GCT TAT GGG CTG ACC TTG CAA CGA GCT GAC CCT	1602
Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Asp Pro	
490 495 500	
TTG GTA GTG CAC CCG AAG CCT AGG TAGGCACCTC ATGTTTATCA AACTTAGGAC	1656

Leu Val Val His Pro Lys Pro Arg

510

1711

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Thr Ser Leu Leu Val Pro Cys Thr Phe Ile Leu Gly Phe Leu Leu Leu
20 25 30

Tyr Ser Phe Leu Asn Lys Lys Val Lys Pro Leu Pro Pro Gly Pro Lys
35 40 45

Pro Trp Pro Ile Val Gly Asn Leu Pro His Leu Gly Pro Lys Pro His
50 55 60

Gln Ser Met Ala Ala Leu Ala Arg Val His Gly Pro Leu Ile His Leu
65 70 75 80

Lys Met Gly Phe Val His Val Val Val Ala Ser Ser Ala Ser Val Ala
85 90 95

Glu Lys Phe Leu Lys Val His Asp Ala Asn Phe Ser Ser Arg Pro Pro
100 105 110

- 130 -

Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr Gln Asp Leu Val Phe
 115 120 125

Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu Arg Lys Ile Cys Ala Leu
 130 135 140

His Leu Phe Ser Ala Lys Ala Leu Asn Asp Phe Thr His Val Arg Gln
 145 150 155 160

Asp Glu Val Gly Ile Leu Thr Arg Val Leu Ala Asp Ala Gly Glu Thr
 165 170 175

Pro Leu Lys Leu Gly Gln Met Met Asn Thr Cys Ala Thr Asn Ala Ile
 180 185 190

Ala Arg Val Met Leu Gly Arg Arg Val Val Gly His Ala Asp Ser Lys
 195 200 205

Ala Glu Glu Phe Lys Ala Met Val Val Glu Leu Met Val Leu Ala Gly
 210 215 220

Val Phe Asn Leu Gly Asp Phe Ile Pro Pro Leu Glu Lys Leu Asp Leu
 225 230 235 240

Gln Gly Val Ile Ala Lys Met Lys Lys Leu His Leu Arg Phe Asp Ser
 245 250 255

Phe Leu Ser Lys Ile Leu Gly Asp His Lys Ile Asn Ser Ser Asp Glu
 260 265 270

Thr Lys Gly His Ser Asp Leu Leu Asn Met Leu Ile Ser Leu Lys Asp
 275 280 285

Ala Asp Asp Ala Glu Gly Gly Arg Leu Thr Asp Val Glu Ile Lys Ala
 290 295 300

Leu Leu Leu Asn Leu Phe Ala Ala Gly Thr Asp Thr Thr Ser Ser Thr

115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300

- 131 -

305	310	315	320
Val Glu Trp Cys Ile Ala Glu Leu Val Arg His Pro Glu Ile Leu Ala			
325	330	335	
Gln Val Gln Lys Glu Leu Asp Ser Val Val Gly Lys Asn Arg Val Val			
340	345	350	
Lys Glu Ala Asp Leu Ala Gly Leu Pro Phe Leu Gln Ala Val Val Lys			
355	360	365	
Glu Asn Phe Arg Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Ile			
370	375	380	
Ala His Glu Ser Cys Glu Val Asn Gly Tyr Leu Ile Pro Lys Gly Ser			
385	390	395	400
Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp Pro Asn Val Trp			
405	410	415	
Asp Glu Pro Leu Glu Phe Arg Pro Glu Arg Phe Leu Lys Gly Gly Glu			
420	425	430	
Lys Pro Asn Val Asp Val Arg Gly Asn Asp Phe Glu Leu Ile Pro Phe			
435	440	445	
Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu Gly Ile Arg Met			
450	455	460	
Val Gln Leu Leu Thr Ala Thr Leu Asn His Ala Phe Asp Phe Asp Leu			
465	470	475	480
Ala Asp Gly Gln Leu Pro Glu Ser Leu Asn Met Glu Glu Ala Tyr Gly			
485	490	495	
Leu Thr Leu Gln Arg Ala Asp Pro Leu Val Val His Pro Lys Pro Arg			
500	505	510	

- 132 -

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 971 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..811

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAT ATG CTT AGC ACT TTA ATC TCC CTT AAA GGA ACT GAT CTT GAC GGT	48
Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly Thr Asp Leu Asp Gly	
1 5 10 15	
GAC GGA GGA AGC TTA ACG GAT ACT GAG ATT AAA GCC TTG CTA TTG AAC	96
Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn	
20 25 30	
ATG TTC ACA GCT GGA ACT GAC ACG TCA GCA AGT ACG GTG GAC TGG GCT	144
Met Phe Thr Ala Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp Ala	
35 40 45	
ATA GCT GAA CTT ATC CGT CAC CCG GAT ATA ATG GTT AAA GCC CAA GAA	192
Ile Ala Glu Leu Ile Arg His Pro Asp Ile Met Val Lys Ala Gln Glu	
50 55 60	
GAA CTT GAT ATT GTT GTG GGC CGT GAC AGG CCT GTT AAT GAA TCA GAC	240
Glu Leu Asp Ile Val Val Gly Arg Asp Arg Pro Val Asn Glu S r Asp	
65 70 75 80	

- 133 -

ATC GCT CAG CTT CCT TAC CTT CAG GCG GTT ATC AAA GAG AAT TTC AGG	288
Ile Ala Gln Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Asn Phe Arg	
85 90 95	
CTT CAT CCA CCA ACA CCA CTC TCG TTA CCA CAC ATC GCG TCA GAG AGC	336
Leu His Pro Pro Thr Pro Leu Ser Leu Pro His Ile Ala Ser Glu Ser	
100 105 110	
TGT GAG ATC AAC GGC TAC CAT ATC CCG AAA GGA TCG ACT CTA TTT GAC	384
Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr Leu Phe Asp	
115 120 125	
GGA CAT ATG GGC CTA GGC CGT GAC CCG GAT CAA TGG TCC GAC CCG TTA	432
Gly His Met Gly Leu Gly Arg Asp Pro Asp Gln Trp Ser Asp Pro Leu	
130 135 140	
GCA TTT AAA CCC GAG AGA TTC TTA CCC GGT GGT GAA AAA TCC GGC GTT	480
Ala Phe Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Ser Gly Val	
145 150 155 160	
GAT GTG AAA GGA AGC GAT TTC GAG CTA ATA CCG TTC GGG GCT GGG AGG	528
Asp Val Lys Gly Ser Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg	
165 170 175	
CCA ATC TGT GCA GGT TTA AGT TTA GGG CTA CGT ACA GAT TTA AGT TGC	576
Pro Ile Cys Ala Gly Leu Ser Leu Gly Leu Arg Thr Asp Leu Ser Cys	
180 185 190	
CTT CAC GCC AAC GTT GCT CAC AAG CAT TTG ATT GGG AAC TTC AGC TGG	624
Leu His Ala Asn Val Ala His Lys His Leu Ile Gly Asn Phe Ser Trp	
195 200 205	
AGA AGT TAC GCC GGA CAA CCT GAA TAT CGC AGG AAA AGT TTA CTG GGC	672
Arg Ser Tyr Ala Gly Gln Pro Glu Tyr Arg Arg Lys Ser Leu Leu Gly	
210 215 220	
TTT ACA CTG CAA AGA GCG GTT CCT TCG GTG GTA CAC CCT AAG CCA AGG	720

- 134 -

Phe Thr Leu Gln Arg Ala Val Pro Ser Val Val His Pro Lys Pro Arg
 225 230 235 240

TTG GCC CCG AAC GTT TAT GGA CCC CGG GTC GGC TTA AAA TTT AAC TTT 768
 Leu Ala Pro Asn Val Tyr Gly Pro Arg Val Gly Leu Lys Phe Asn Phe
 245 250 255

GCT TCT TGG ACA AGG TAT ATG GCT TGC ACG AAA CTA ACG TTT T 811
 Ala Ser Trp Thr Arg Tyr Met Ala Cys Thr Lys Leu Thr Phe
 260 265 270

AACACACCGT AGTTTGATCC GGAGTTAGCT TTATGTAAGA ACGTGTAACG CCAAATCAAG 871

CCATTATCAA CTACCGTGAG CTGTTTGTAC CCTATCTATA AATCTTGAAG AGGAACATTT 931

CAGAACTCTT GACTATGTTT CAGGAACAAA AAAAAAAAAA 971

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 270 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly Thr Asp Leu Asp Gly
 1 5 10 15

Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn
 20 25 30

Met Phe Thr Ala Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp Ala
 35 40 45

- 135 -

Ile Ala Glu Leu Ile Arg His Pro Asp Ile Met Val Lys Ala Gln Glu
50 55 60

- 136 -

245	250	255
Ala Ser Trp Thr Arg Tyr Met Ala Cys Thr Lys Leu Thr Phe		
260	265	270

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6595 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1478..1927

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2651..3091

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3170..3340

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3421..3900

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GTCGACTCTC TCCCTTTCGC TTGCTACTTT TTCTACATAA ATAAATGCAA TGATAAATTT

60

- 137 -

GTGCACACAT TCGTATGTTT GAAACATGGT AGGATCCACA ATTTATACTT TATAGACTCA	120
AAATGGAAAA GAAACGTACA TTATAAATTT ATCTGCAATT TGTTTTCTCT TGCTAAACTA	180
GACTGTATAA TAACCTCTGT ATATGCTATT ACTCGATTGT AAACGTACCC CGCAAGTCGC	240
AAGCAAGGTA AATAAAGTAT AATTATATTT TCACACACGA AACTTTAATT ATTATTTTTA	300
TCACTTGCAG ATTAACAGTA AAAAAAAAAA AAATGTGACT TTAACGGCGA CAAAACTAC	360
TGATCTTTCT CCAATATTTA AATAATATAA TTAATAAACG TCTTTTCATA CTTGTATTTT	420
CCGACCCGAG TTCTGAAAGT GAAAACATAT GGTACTAGAT ATTCTCGATT TGTTTTGTAG	480
CCACTAGACT CTAAACAGAA AAAAGAAGCC AAAAGGACAA CGTTAAAAAA GAGACACTGT	540
TATTAAAAGT TAGAAACCAA ACGGTGAAAA TCCAGCTACA TACATAAAAT AAAGCCAAGG	600
TACCAAACTA ATGAACTGTA ACCTCTTTTT TCTTTTCTTT TTTGTTAAAG GATTTATGAA	660
CTGTAACTTA GAATGCTTGG TTTGTGGGCA GTGTAATATA TGACACACAT GCATTTTTTT	720
TGTTTGTCAA ATAGGAAGAC TTCTTTTTTC TTTATCAACT TCCTTATTTT CATAAAACAA	780
AACACTGAAA AAAGTACAGA TGTTCTCACG TACGTCACGT GTACATACAT ATATATTAGA	840
CCACTATATA ATAAGATATG AAGTGTTAGG TTTAAATCAA TTAACGAATC CCATCCAAAT	900
GATGAAACAG TTAACAAGAA ATCAAAATAG TTTATTAGGG TTACAATGAT TTTATACTTT	960
TAAGAAATCT TAGAACCTAT CACTTACAAA TGAGTAAATG ACCATTACTC CTCGAGAATC	1020
TAAGGCGCTT AAGGAAGCAT TGCGAATCGG GTGTGAAAAA GATCTATTTT TTGAATTATT	1080
TCACACAATT TCTTAATGTC AATTTTCGAT GCTCCCATAT TCTCCACGGT TTAAAGCAAG	1140
ATTGGTGGGA AAGGGATATT CTCGCATCGA TTACAATGAA ATATGGGTTG AAAAAAAAAA	1200

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

- 138 -

AAAAAATTA CTCAATGTTG CACCAAAAAC CAGAAAAC TC TAAGTTGCGC TAATAAAAAA	1260
AAAAGTTATA AACCCAACAT CAAACCAAAA CCGTACTAAA CTGTCCCAT A TGAGATTAG	1320
CTTTAAATAA ATTAGTACTT CTCATAACGA TAACTAAATT AAATTTCCCT AGCCAAGACA	1380
TACATATAGT TTTGATTGAC AAAAAAAAAA AAAACTCCTC TATTTATAGC TTGTGTTTGT	1440
TTTCCTCATT TTCACTTAC CATTCAAACC CAACACT ATG GCA ACT CTA TTT CTC	1495
Met Ala Thr Leu Phe Leu	
1 5	
ACA ATC CTC CTA GCC ACT GTC CTC TTC CTC ATC CTC CGT ATC TTC TCT	1543
Thr Ile Leu Leu Ala Thr Val Leu Phe Leu Ile Leu Arg Ile Phe Ser	
10 15 20	
CAC CGT CGC AAC CGC AGC CAC AAC AAC CGT CTT CCA CCG GGG CCA AAC	1591
His Arg Arg Asn Arg Ser His Asn Asn Arg Leu Pro Pro Gly Pro Asn	
25 30 35	
CCA TGG CCC ATC ATC GGA AAC CTC CCT CAC ATG GGC ACT AAG CCT CAT	1639
Pro Trp Pro Ile Ile Gly Asn Leu Pro His Met Gly Thr Lys Pro His	
40 45 50	
CGA ACC CTT TCC GCC ATG GTT ACT ACT TAC GGC CCT ATC CTC CAC CTC	1687
Arg Thr Leu Ser Ala Met Val Thr Thr Tyr Gly Pro Ile Leu His Leu	
55 60 65 70	
CGA CTA GGG TTC GTA GAC GTC GTG GTC GCC GCT TCT AAA TCC GTG GCC	1735
Arg Leu Gly Phe Val Asp Val Val Val Ala Ala Ser Lys Ser Val Ala	
75 80 85	
GAG CAG TTC TTG AAA ATA CAC GAC GCC AAT TTC GCT AGC CGA CCA CCA	1783
Glu Gln Phe Leu Lys Ile His Asp Ala Asn Phe Ala Ser Arg Pro Pro	
90 95 100	
AAC TCA GGA GCC AAA CAC ATG GCA TAT AAC TAT CAA GAT CTT GTC TTT	1831

- 139 -

Asn Ser Gly Ala Lys His Met Ala Tyr Asn Tyr Gln Asp Leu Val Phe

105

110

115

GCA CCT TAC GGA CAC CGA TGG AGA CTG TTG AGA AAG ATT AGT TCT GTT 1879

Ala Pro Tyr Gly His Arg Trp Arg Leu Leu Arg Lys Ile Ser Ser Val

120

125

130

CAT CTA TTT TCA GCT AAA GCT CTC GAA GAT TTC AAA CAT GTT CGA CAG 1927

His Leu Phe Ser Ala Lys Ala Leu Glu Asp Phe Lys His Val Arg Gln

135

140

145

150

GTAAACAAT TATAACGGT ATTCTCATTT TCTAACGCTA TAGCTCACTG GCCTGTAATC 1987

ATGTCATTTT AATGTTTTGA CTTTTTCTTT ATATATACAT AATTATAATT TATAATTGGG 2047

ATTTCAAACC CTATCTCTCA CTATTTCAAG ACTAGACCGG ATTGGAATTT GAACTTTTGT 2107

AATGAATATT AGTATCTGCA CATAAATTTT ATGTTAAAGT TGGGTTTTCT TAAAGTGAAT 2167

TTATATATTA AAAATATATA AACGATTGGG TTTTACTCAA ATGAATTTAC ATAAGAGCTA 2227

GGTATAAGTG CAAATATGCA ATACTGTCAT TGTCGTGGAT GTATAAAAGT ATGATCTAAC 2287

TTTGATGATG CCATGGAAAA ATTGGAAAGT TCAGATCCAG AGGAAACATT GCTTGAATTA 2347

TAAAATGTAT GGACCACATT GTTTCCTTAA ATGGAAGGTC TCACGAGTTT CTCAATTTCA 2407

GACTACTGAT AATATATGCT ATTATAGATT TTATTTTCTG ATTATTTTTT TTGGTTTAAT 2467

TTAATTAGAG TAAATTTTAA AAAAGAAATA TATGGTTTTG TTAACCGTGT TTTAAAATTT 2527

GATAGAGCTT TTAGATCATA ATCATAATTT TTTCGTATTA ATTGTGATTA TGTTGGACGA 2587

AAATACTTAA TTAGTATTCA AGAAAACTCT TATTCTAAAA ACAGAAATAA ATGAATTTTA 2647

CAG GAA GAG GTT GGA ACG CTA ACG CGG GAG CTA GTG CGT GTT GGC ACG 2695

Glu Glu Val Gly Thr Leu Thr Arg Glu Leu Val Arg Val Gly Thr

- 140 -

1	5	10	15	
AAA CCC GTG AAT TTA GGC CAG TTG GTG AAC ATG TGT GTA GTC AAC GCT				2743
Lys Pro Val Asn Leu Gly Gln Leu Val Asn Met Cys Val Val Asn Ala				
20	25	30		
CTA GGA CGA GAG ATG ATC GGA CGG CGA CTG TTC GGC GCC GAC GCC GAT				2791
Leu Gly Arg Glu Met Ile Gly Arg Arg Leu Phe Gly Ala Asp Ala Asp				
35	40	45		
CAT AAA GCT GAC GAG TTT CGA TCG ATG GTG ACG GAA ATG ATG GCT CTC				2839
His Lys Ala Asp Glu Phe Arg Ser Met Val Thr Glu Met Met Ala Leu				
50	55	60		
GCC GGA GTA TTT AAC ATC GGA GAT TTC GTG CCG TCA CTT GAT TGG TTA				2887
Ala Gly Val Phe Asn Ile Gly Asp Phe Val Pro Ser Leu Asp Trp Leu				
65	70	75		
GAT TTA CAA GGC GTC GCT GGT AAA ATG AAA CGG CTT CAC AAA AGA TTC				2935
Asp Leu Gln Gly Val Ala Gly Lys Met Lys Arg Leu His Lys Arg Phe				
80	85	90	95	
GAC GCT TTT CTA TCG TCG ATT TTG AAA GAG CAC GAA ATG AAC GGT CAA				2983
Asp Ala Phe Leu Ser Ser Ile Leu Lys Glu His Glu Met Asn Gly Gln				
100	105	110		
GAT CAA AAG CAT ACA GAT ATG CTT AGC ACT TTA ATC TCC CTT AAA GGA				3031
Asp Gln Lys His Thr Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly				
115	120	125		
ACT GAT CTT GAC GGT GAC GGA GGA AGC TTA ACG GAT ACT GAG ATT AAA				3079
Thr Asp Leu Asp Gly Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys				
130	135	140		
GCC TTG CTA TTG GTCAGTTTTT TGACAATTAA TTTCCTTAAA AATCGTATAT				3131
Ala Leu Leu Leu				
145				

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963

- 141 -

AATGAAAGTT AGATTGTTTT TTTTGGTTGT AAATACAG AAC ATG TTC ACA GCT	3184
Asn Met Phe Thr Ala	
1 5	
GGA ACT GAC ACG TCA GCA AGT ACG GTG GAC TGC GCT ATA GCT GAA CTT	3232
Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp Ala Ile Ala Glu Leu	
10 15 20	
ATC CGT CAC CCG GAT ATA ATG GTT AAA GCC CAA GAA GAA CTT GAT ATT	3280
Ile Arg His Pro Asp Ile Met Val Lys Ala Gln Glu Glu Leu Asp Ile	
25 30 35	
GTT GTG GGC CGT GAC AGG CCT GTT AAT GAA TCA GAC ATC GCT CAG CTT	3328
Val Val Gly Arg Asp Arg Pro Val Asn Glu Ser Asp Ile Ala Gln Leu	
40 45 50	
CCT TAC CTT CAG GTACCGTTAA CCCAAACCGG AATTTGGAAT TGTTTTGGTT	3380
Pro Tyr Leu Gln	
55	
AGCGAGCTAT TGTTGTTAAT CCGGTTTTGG TTTAAAACAG GCG GTT ATC AAA GAG	3435
Ala Val Ile Lys Glu	
1 5	
AAT TTC AGG CTT CAT CCA CCA ACA CCA CTC TCG TTA CCA CAC ATC GCG	3483
Asn Phe Arg Leu His Pro Pro Thr Pro Leu Ser Leu Pro His Ile Ala	
10 15 20	
TCA GAG AGC TGT GAG ATC AAC GGC TAC CAT ATC CCG AAA GGA TCG ACT	3531
Ser Glu Ser Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr	
25 30 35	
CTA TTG ACG AAC ATA TGG GCC ATA GCC CGT GAC CCG GAT CAA TGG TCC	3579
Leu Leu Thr Asn Ile Trp Ala Ile Ala Arg Asp Pro Asp Gln Trp Ser	
40 45 50	
GAC CCG TTA GCA TTT AAA CCC GAG AGA TTC TTA CCC GGT GGT GAA AAA	3627

- 142 -

Asp Pro Leu Ala Phe Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys	
55 60 65	
TCC GGC GTT GAT GTG AAA GGA AGC GAT TTC GAG CTA ATA CCG TTC GGA	3675
Ser Gly Val Asp Val Lys Gly Ser Asp Phe Glu Leu Ile Pro Phe Gly	
70 75 80 85	
GCT GGG AGG AGA ATC TGT GCC GGT TTA AGT TTA GGG TTA CGT ACG ATT	3723
Ala Gly Arg Arg Ile Cys Ala Gly Leu Ser Leu Gly Leu Arg Thr Ile	
90 95 100	
CAG TTT CTT ACG GCG ACG TTG GTT CAA GGA TTT GAT TGG GAA TTA GCT	3771
Gln Phe Leu Thr Ala Thr Leu Val Gln Gly Phe Asp Trp Glu Leu Ala	
105 110 115	
GGA GGA GTT ACG CCG GAG AAG CTG AAT ATG GAG GAG AGT TAT GGG CTT	3819
Gly Gly Val Thr Pro Glu Lys Leu Asn Met Glu Glu Ser Tyr Gly Leu	
120 125 130	
ACA CTG CAA AGA GCG GTT CCT TTG GTG GTA CAT CCT AAG CCA AGG TTG	3867
Thr Leu Gln Arg Ala Val Pro Leu Val Val His Pro Lys Pro Arg Leu	
135 140 145	
GCT CCG AAC GTT TAT GGA CTC GGG TCG GGT TAAAATTTAA CTTTGCTTCT	3917
Ala Pro Asn Val Tyr Gly Leu Gly Ser Gly	
150 155 160	
TGGACAAGGT ATATGGCTTG CACGAAAATA AAGTTTTTAAA ACAGCGTAGT TTGATCCGGA	3977
GTTAGCTTTA TGTAAGAACG TGTAACGCCA AATCAAGTCA TTATTAAATA TTGTGAGTTG	4037
TTTGTAACCT ATATATAAAT CTTGAAGAGG AAGATTTTCTAG AAATCTTGAA TATGTTTTAG	4097
GAAAAACATT GTTTTTTTTTA CAGTAGCGCA AGTTGAATTA AAACCTATTC CTTACAGAAC	4157
CAAATGCATT AATAATTCTA GATATTTTTTG GCCAAGACAA TCAGATTTTT CAATATTTCA	4217

- 143 -

TATATACTAG GTGGAACACC ACCACCTGCA ACTCTGCAAC ACATGTTACG TTACACAATC	4277
ACTTTTGGCG GTTTTC AATT ATTTATATAA AATTGTAAAT GTTTGTACAC AGTAGAAAAT	4337
TAGTAATAGT GAATTTTGT TCTCCGAATA TGTATAGCAA TATATATGGC ATGGATCAAA	4397
CTAGCCGACA TCCTAACTTG TTCACAGCTT TCCTTTTTAC TTATCTAGTC GATTAAGCAT	4457
CAGAAAGTAT GTTTTAATTT TTAAATTTGA AAAAGGTGTA CTTACAAGTT CGGGTGTTC	4517
CACGGAGGAG AGCTACAATA ATGAAAAGC TGACTCAAGA AGGGCTATAG AAGAAACAAG	4577
AGTCACGGAA CAAGTTGTCA CTCTCAATCT CCAGTACACT AGCTTCCATA ACTCTCTCTC	4637
TTTCTCTCTT TCTTCTCTCT CTAAAAGTTA TCAGAATAGA AATCTCTCTC TCTCAACAAG	4697
TCTAACAGTG CCATTTGTAT CTCTGAACTC CAACATGGCT CCTCTGGTTC TCTACCTTCT	4757
CACTCTCCTC ATGGCTGGCC ATTCCAGTAA GAACTCTCAC TGATCTTCTT CACCTTTGTT	4817
TATGGATTTG GTCTCTCAGT CTCACTCTCG CTTACCCTTT CACATTCAGC TCTGGCTCTC	4877
TGGTTTAAGA AACCTTAAT CTACAAAGCT TGCTTTCCTC GCAAATGAAC TACCTTACTT	4937
ATCTCTTATG CAACTCTTGT TGATGATTTG CAAACATCTT AACCTCTCGA AACAAGATTT	4997
ACAAATCTTA CTGGCTTCAC TTACAATTTT GTTCCCATTT TTTTCTTCTT TGGTAGGTGC	5057
CTCATGGTGT GTGTGCAAAA CAGGGCTGAG TGACTCAGTG CTACAAAAGA CATTAGACTA	5117
TGCTTGTTGA AATGGAGCTG ACTGTAACCC AACTCACCCA AAAGGCTCTT GCTTCAATCC	5177
TGACAATGTT AGGGCTCATT GCAACTATGC AGTCAATAGC TTCTTCCAAA AGAAAGGTCA	5237
AGCTTCTGAG TCTTGTAAC TCACTGGTAC TGCCACTCTT ACCACCACCG ATCCCAGTAA	5297
GTTTTCAGAA TGTTAACT CTTGTGATCT TTAGAACCCT ACAAATTTT GAGTCTCAGA	5357

- 144 -

AAGTTCAAGT TCAAGGTCTT TTGGTTAGAG TACTAAAGAT TCAAGTAGAG ACTAGGCGTG 5417

AGATATTTTT TCTCTGATGT GTGATTTTTT GGCACAGGCT ATACAGGATG TGCATTCCCT 5477

TCTAGTGCTA GGTACGGCTC TTTGCTTCTC TACACATTTA TTTTCTTAAT GGCTTTATCT 5537

AGAACTTTGA AGGATACCAT TTTATTTTTT TTGGACAAAG AAGGATAGCC ATTTAATACT 5597

ACACTTTAAT GTTGGATTAA CTAACCTATT ATGCCTATTT AATGGCCTAC ACTTTAAGTG 5657

GACACAAGCT TGATTTGGTT ATAAAAAAG TGCACATAA TCTTATTTTA CTGAACCCCTT 5717

TTTTCTATGA TTTTTTTACT AAACCTTAGA TAACATCTAC AACAATTCAA TTGCCTTTTT 5777

TTGGGGATTG TATAAGTTTG AACCTATGGT TAGTGTATTG ACTTGCGCGT CTCTTATTGC 5837

AACGGTTCTT TGAAAACACA TTAATGATAA ATAAATTGAA AAGTATAGAG ATGGCAATTG 5897

TTTCAAAAGC TAATCTTTCT GCTTGCTAAT ACTTTACATA AAAACAAAAA AATTAAGAAG 5957

ATTTTCAAAC AATACAACCT TTTTACCTTG TCCTAACAAA TTCAACTCAA ATGACATGTG 6017

TTTGCTTTAA AATAGTAACA ACTGTAAATT CATTGCTCT TGAGACATAA GTGCAAGCTA 6077

AAGATAAAG CAAGCAATAC AATTAGGCCT AATTAAGATT ACGAATATTG TTGTTTGTTT 6137

ATAGTGTTTC TAGTGGAAGC GGTAGCACCA CCGTGACGCC AGGCAAAAAC AGTCCAAAAG 6197

GAAGCAACAG CATCACCACA TTTCCCGGCG GAAACAGTCC ATACACTGGC ACACCATCCA 6257

CCGGATTATT AGGAGGCAAT ATCACTGATG CAACTGGAAC CGGGTTGAAC CCGGATTACT 6317

CAACCGAAAG CAGTGGATTT GCGCTCTATT ACTCCAACAA CCTTCTGTTA ACCGGCTTTT 6377

GTTCTCTCGT GATGATGCTC TGAAGAAGAA TCACCGTCTT CTTTTAGTTT ATGCTTAGTC 6437

AAAAAATAT GTTATTTATA TGTCTTGTT GTTTTAGAGA TAATTTAATC TGGATTTCGG 6497

- 145 -

TTCTTTTTTA CTTTCCGGTT TTAAGAAAAC AATTATCAAT GTAAAACCAA ATCTACTATC 6557

GATCGGTTTG GTACGAATTC CTGCAGCCCG GGGGATCC 6595

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 149 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Thr Leu Phe Leu Thr Ile Leu Leu Ala Thr Val Leu Phe Leu
1 5 10 15

Ile Leu Arg Ile Phe Ser His Arg Arg Asn Arg Ser His Asn Asn Arg
20 25 30

Leu Pro Pro Gly Pro Asn Pro Trp Pro Ile Ile Gly Asn Leu Pro His
35 40 45

Met Gly Thr Lys Pro His Arg Thr Leu Ser Ala Met Val Thr Thr Tyr
50 55 60

Gly Pro Ile Leu His Leu Arg Leu Gly Phe Val Asp Val Val Val Ala
65 70 75 80

Ala Ser Lys Ser Val Ala Glu Gln Phe Leu Lys Ile His Asp Ala Asn
85 90 95

Phe Ala Ser Arg Pro Pro Asn Ser Gly Ala Lys His Met Ala Tyr Asn
100 105 110

- 146 -

Tyr Gln Asp Leu Val Phe Ala Pro Tyr Gly His Arg Trp Arg Leu Leu
 115 120 125

Arg Lys Ile Ser Ser Val His Leu Phe Ser Ala Lys Ala Leu Glu Asp
 130 135 140

Phe Lys His Val Arg Gln
 145 150

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Glu Glu Val Gly Thr Leu Thr Arg Glu Leu Val Arg Val Gly Thr Lys
 1 5 10 15

Pro Val Asn Leu Gly Gln Leu Val Asn Met Cys Val Val Asn Ala Leu
 20 25 30

Gly Arg Glu Met Ile Gly Arg Arg Leu Phe Gly Ala Asp Ala Asp His
 35 40 45

Lys Ala Asp Glu Phe Arg Ser Met Val Thr Glu Met Met Ala Leu Ala
 50 55 60

Gly Val Phe Asn Ile Gly Asp Phe Val Pro Ser Leu Asp Trp Leu Asp
 65 70 75 80

Leu Gln Gly Val Ala Gly Lys Met Lys Arg Leu His Lys Arg Phe Asp

- 147 -

85

90

95

Ala Phe Leu Ser Ser Ile Leu Lys Glu His Glu Met Asn Gly Gln Asp

100

105

110

Gln Lys His Thr Asp Met Leu Ser Thr Leu Ile Ser Leu Lys Gly Thr

115

120

125

Asp Leu Asp Gly Asp Gly Gly Ser Leu Thr Asp Thr Glu Ile Lys Ala

130

135

140

Leu Leu Leu

145

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Asn Met Phe Thr Ala Gly Thr Asp Thr Ser Ala Ser Thr Val Asp Trp

1

5

10

15

Ala Ile Ala Glu Leu Ile Arg His Pro Asp Ile Met Val Lys Ala Gln

20

25

30

Glu Glu Leu Asp Ile Val Val Gly Arg Asp Arg Pro Val Asn Glu Ser

35

40

45

Asp Ile Ala Gln Leu Pro Tyr Leu Gln

50

55

- 148 -

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 160 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ala	Val	Ile	Lys	Glu	Asn	Phe	Arg	Leu	His	Pro	Pro	Thr	Pro	Leu	Ser
1				5					10					15	
Leu	Pro	His	Ile	Ala	Ser	Glu	Ser	Cys	Glu	Ile	Asn	Gly	Tyr	His	Ile
			20					25					30		
Pro	Lys	Gly	Ser	Thr	Leu	Leu	Thr	Asn	Ile	Trp	Ala	Ile	Ala	Arg	Asp
			35					40					45		
Pro	Asp	Gln	Trp	Ser	Asp	Pro	Leu	Ala	Phe	Lys	Pro	Glu	Arg	Phe	Leu
		50				55					60				
Pro	Gly	Gly	Glu	Lys	Ser	Gly	Val	Asp	Val	Lys	Gly	Ser	Asp	Phe	Glu
	65				70					75				80	
Leu	Ile	Pro	Phe	Gly	Ala	Gly	Arg	Arg	Ile	Cys	Ala	Gly	Leu	Ser	Leu
			85						90					95	
Gly	Leu	Arg	Thr	Ile	Gln	Phe	Leu	Thr	Ala	Thr	Leu	Val	Gln	Gly	Phe
			100					105						110	
Asp	Trp	Glu	Leu	Ala	Gly	Gly	Val	Thr	Pro	Glu	Lys	Leu	Asn	Met	Glu
			115				120							125	
Glu	Ser	Tyr	Gly	Leu	Thr	Leu	Gln	Arg	Ala	Val	Pro	Leu	Val	Val	His

- 149 -

130 135 140

Pro Lys Pro Arg Leu Ala Pro Asn Val Tyr Gly Leu Gly Ser Gly

145 150 155

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1748 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 22..1563

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGTCGAGAAA GAAGAACAGC C ATG TTT CTC ATA GTA GTG ATC ACC TTC CTC	51
Met Phe Leu Ile Val Val Ile Thr Phe Leu	
1 5 10	
TTC GCC GTG TTT TTG TTC CGG CTT CTT TTC TCC GGC AAA TCC CAA CGC	99
Phe Ala Val Phe Leu Phe Arg Leu Leu Phe Ser Gly Lys Ser Gln Arg	
15 20 25	
CAC TCG CTC CCT CTC CCT CCT GGC CCC AAA CCA TGG CCG GTG GTT GGC	147
His Ser Leu Pro Leu Pro Pro Gly Pro Lys Pro Trp Pro Val Val Gly	
30 35 40	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070

- 150 -

AAC TTA CCT CAC TTG GGC CCC TTC CCG CAC CAC TCC ATC GCG GAG TTG	195
Asn Leu Pro His Leu Gly Pro Phe Pro His His Ser Ile Ala Glu Leu	
45 50 55	
GCG AAG AAA CAC GGG CCG CTC ATG CAC CTC CGC CTC GGC TAC GTT GAC	243
Ala Lys Lys His Gly Pro Leu Met His Leu Arg Leu Gly Tyr Val Asp	
60 65 70	
GTA GTC GTG GCG GCA TCA GCA TCC GTA GCG GCC CAG TTC TTG AAG ACT	291
Val Val Val Ala Ala Ser Ala Ser Val Ala Ala Gln Phe Leu Lys Thr	
75 80 85 90	
CAC GAC GCC AAT TTC TCC AGC CGA CCG CCC AAC TCC GGC GCC AAG CAC	339
His Asp Ala Asn Phe Ser Ser Arg Pro Pro Asn Ser Gly Ala Lys His	
95 100 105	
CTC GCC TAT AAC TAC CAG GAC CTC GTG TTC AGG CCG TAC GGT CCA CGG	387
Leu Ala Tyr Asn Tyr Gln Asp Leu Val Phe Arg Pro Tyr Gly Pro Arg	
110 115 120	
TGG CGC ATG TTC CGG AAG ATC AGC TCC GTC CAT CTG TTC TCC GGC AAA	435
Trp Arg Met Phe Arg Lys Ile Ser Ser Val His Leu Phe Ser Gly Lys	
125 130 135	
GCC TTG GAT GAT CTT AAA CAC GTC CGG CAG GAG GAG GTA AGT GTG CTA	483
Ala Leu Asp Asp Leu Lys His Val Arg Gln Glu Glu Val Ser Val Leu	
140 145 150	
GCG CAT GCC TTG GCA AAT TCA GGG TCA AAG GTA GTG AAC CTG GCG CAA	531
Ala His Ala Leu Ala Asn Ser Gly Ser Lys Val Val Asn Leu Ala Gln	
155 160 165 170	
CTG CTG AAC CTG TGC ACG GTC AAT GCT CTA GGA AGG GTG ATG GTA GGG	579
Leu Leu Asn Leu Cys Thr Val Asn Ala Leu Gly Arg Val Met Val Gly	
175 180 185	
CGG AGG GTT TTC GGC GAC GGC AGC GGA GGC GAC GAT CCG AAG GCG GAC	627

- 151 -

Arg Arg Val Phe Gly Asp Gly Ser Gly Gly Asp Asp Pro Lys Ala Asp
 190 195 200

GAG TTC AAA TCG ATG GTG GTG GAG ATG ATG GTG TTG GCA GGA GTG TTC 675
 Glu Phe Lys Ser Met Val Val Glu Met Met Val Leu Ala Gly Val Phe
 205 210 215

AAC ATA GGT GAC TTC ATC CCC TCT CTC GAA TGG CTT GAC TTG CAA GGC 723
 Asn Ile Gly Asp Phe Ile Pro Ser Leu Glu Trp Leu Asp Leu Gln Gly
 220 225 230

GTG GCG TCC AAG ATG AAG AAG CTC CAC AAG AGA TTC GAC GAC TTC TTG 771
 Val Ala Ser Lys Met Lys Lys Leu His Lys Arg Phe Asp Asp Phe Leu
 235 240 245 250

ACA GCC ATT GTC GAG GAC CAC AAG AAG GGC TCC GGC ACG GCG GGG CAC 819
 Thr Ala Ile Val Glu Asp His Lys Lys Gly Ser Gly Thr Ala Gly His
 255 260 265

GTC GAC ATG TTG ACC ACT CTG CTC TCG CTC AAG GAA GAC GCC GAC GGC 867
 Val Asp Met Leu Thr Thr Leu Leu Ser Leu Lys Glu Asp Ala Asp Gly
 270 275 280

GAA GGA GGC AAG CTC ACC GAT ACT GAA ATC AAA GCT TTG CTT TTG AAC 915
 Glu Gly Gly Lys Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn
 285 290 295

ATG TTC ACG GCT GGC ACT GAT ACG TCA TCG AGC ACG GTG GAA TGG GCA 963
 Met Phe Thr Ala Gly Thr Asp Thr Ser Ser Ser Thr Val Glu Trp Ala
 300 305 310

ATA GCT GAA CTC ATT CGG CAC CCT CAT ATG CTA GCG CGA GTT CAG AAA 1011
 Ile Ala Glu Leu Ile Arg His Pro His Met Leu Ala Arg Val Gln Lys
 315 320 325 330

GAG CTT GAC GAT TTT GTT GGC CAT GAC CGA CTT GTG ACC GAA TCC GAC 1059
 Glu Leu Asp Asp Phe Val Gly His Asp Arg Leu Val Thr Glu Ser Asp

- 152 -

335	340	345	
ATA CCC AAC CTC CCT TAC CTC CAA GCC GTG ATC AAG GAA ACG TTC CGA			1107
Ile Pro Asn Leu Pro Tyr Leu Gln Ala Val Ile Lys Glu Thr Phe Arg			
350	355	360	
CTC CAC CCA TCC ACT CCT CTC TCG TTG CCT CGT ATG GCA GCC GAG AGT			1155
Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Met Ala Ala Glu Ser			
365	370	375	
TGC GAA ATC AAC GGG TAC CAC ATC CCG AAA GGC TCC ACA CTC TTG GTC			1203
Cys Glu Ile Asn Gly Tyr His Ile Pro Lys Gly Ser Thr Leu Leu Val			
380	385	390	
AAT GTA TGG GCC ATA TCG CGT GAC CCG GCT GAA TGG GCC GAC CCA CTG			1251
Asn Val Trp Ala Ile Ser Arg Asp Pro Ala Glu Trp Ala Asp Pro Leu			
395	400	405	410
GAG TTC AAG CCC GAG AGG TTC CTG CCG GGG GGC GAA AAG CCT AAT GTT			1299
Glu Phe Lys Pro Glu Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val			
415	420	425	
GAT ATT AGA GGA AAC GAT TTT GAA GTC ATA CCC TTC GGT GCC GGG CGA			1347
Asp Ile Arg Gly Asn Asp Phe Glu Val Ile Pro Phe Gly Ala Gly Arg			
430	435	440	
AGA ATA TGT GCC GGG ATG AGC TTG GGC CTG CGT ATG GTC CAT TTA ATG			1395
Arg Ile Cys Ala Gly Met Ser Leu Gly Leu Arg Met Val His Leu Met			
445	450	455	
ACT GCA ACA TTG GTC CAC GCA TTT AAT TGG GCC TTG GCT GAT GGG CTG			1443
Thr Ala Thr Leu Val His Ala Phe Asn Trp Ala Leu Ala Asp Gly Leu			
460	465	470	
ACC GCT GAG AAG TTA AAC ATG GAT GAA GCA TAT GGG CTC ACT CTA CAA			1491
Thr Ala Glu Lys Leu Asn Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln			
475	480	485	490

Publ. No. WO 97/32023

- 153 -

CGA GCT GCA CCG TTA ATG GTG CAC CCG CGC ACC AGG CTG GCC CCA CAG 1539
 Arg Ala Ala Pro Leu Met Val His Pro Arg Thr Arg Leu Ala Pro Gln
 495 500 505

GCA TAT AAA ACT TCA TCA TCT TAATTAGAGA GCTATGTTCT GGGTGTGCCC 1590
 Ala Tyr Lys Thr Ser Ser Ser
 510

GGTTTGATGT CTCCATGTTT TCTATTTAGG TTAAATCTG TAAGATAAGG TGATTCTATG 1650

CTGAATCACA AAAGTTGCTA TCTAAATTCC ATGTCCAATG AAAACGTTCT TCTTCCCTTC 1710

TTATAATTTA TGAATACTTA TGATATAGGC GACAGCAA 1748

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 513 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Phe Leu Ile Val Val Ile Thr Phe Leu Phe Ala Val Phe Leu Phe
 1 5 10 15
 Arg Leu Leu Phe Ser Gly Lys Ser Gln Arg His Ser Leu Pro Leu Pro
 20 25 30
 Pro Gly Pro Lys Pro Trp Pro Val Val Gly Asn Leu Pro His Leu Gly
 35 40 45
 Pro Phe Pro His His Ser Ile Ala Glu Leu Ala Lys Lys His Gly Pro
 50 55 60

- 154 -

Leu Met His Leu Arg Leu Gly Tyr Val Asp Val Val Val Ala Ala Ser
65 70 75 80

Ala Ser Val Ala Ala Gln Phe Leu Lys Thr His Asp Ala Asn Phe Ser
85 90 95

Ser Arg Pro Pro Asn Ser Gly Ala Lys His Leu Ala Tyr Asn Tyr Gln
100 105 110

Asp Leu Val Phe Arg Pro Tyr Gly Pro Arg Trp Arg Met Phe Arg Lys
115 120 125

Ile Ser Ser Val His Leu Phe Ser Gly Lys Ala Leu Asp Asp Leu Lys
130 135 140

His Val Arg Gln Glu Glu Val Ser Val Leu Ala His Ala Leu Ala Asn
145 150 155 160

Ser Gly Ser Lys Val Val Asn Leu Ala Gln Leu Leu Asn Leu Cys Thr
165 170 175

Val Asn Ala Leu Gly Arg Val Met Val Gly Arg Arg Val Phe Gly Asp
180 185 190

Gly Ser Gly Gly Asp Asp Pro Lys Ala Asp Glu Phe Lys Ser Met Val
195 200 205

Val Glu Met Met Val Leu Ala Gly Val Phe Asn Ile Gly Asp Phe Ile
210 215 220

Pro Ser Leu Glu Trp Leu Asp Leu Gln Gly Val Ala Ser Lys Met Lys
225 230 235 240

Lys Leu His Lys Arg Phe Asp Asp Phe Leu Thr Ala Ile Val Glu Asp
245 250 255

His Lys Lys Gly Ser Gly Thr Ala Gly His Val Asp Met Leu Thr Thr

- 155 -

260	265	270
Leu Leu Ser Leu Lys Glu Asp Ala Asp Gly Glu Gly Gly Lys Leu Thr		
275	280	285
Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Met Phe Thr Ala Gly Thr		
290	295	300
Asp Thr Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg		
305	310	315 320
His Pro His Met Leu Ala Arg Val Gln Lys Glu Leu Asp Asp Phe Val		
325	330	335
Gly His Asp Arg Leu Val Thr Glu Ser Asp Ile Pro Asn Leu Pro Tyr		
340	345	350
Leu Gln Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro		
355	360	365
Leu Ser Leu Pro Arg Met Ala Ala Glu Ser Cys Glu Ile Asn Gly Tyr		
370	375	380
His Ile Pro Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ser		
385	390	395 400
Arg Asp Pro Ala Glu Trp Ala Asp Pro Leu Glu Phe Lys Pro Glu Arg		
405	410	415
Phe Leu Pro Gly Gly Glu Lys Pro Asn Val Asp Ile Arg Gly Asn Asp		
420	425	430
Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met		
435	440	445
Ser Leu Gly Leu Arg Met Val His Leu Met Thr Ala Thr Leu Val His		
450	455	460

- 156 -

Ala Phe Asn Trp Ala Leu Ala Asp Gly Leu Thr Ala Glu Lys Leu Asn
 465 470 475 480

Met Asp Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ala Pro Leu Met
 485 490 495

Val His Pro Arg Thr Arg Leu Ala Pro Gln Ala Tyr Lys Thr Ser Ser
 500 505 510

Ser

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1660 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4..1528

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AAA ATG ACC ATT TTA GCT TTC GTA TTT TAC GCC CTC ATC CTC GGG TCA 48
 Met Thr Ile Leu Ala Phe Val Phe Tyr Ala Leu Ile Leu Gly Ser
 1 5 10 15

GTA CTC TAT GTA TTT CTT AAC TTA AGT TCA CGT AAA TCC GCC AGA CTC 96

- 157 -

Val Leu Tyr Val Phe Leu Asn Leu Ser Ser Arg Lys Ser Ala Arg Leu	
20 25 30	
CCA CCC GGG CCA ACA CCA TGG CCT ATA GTC GGG AAC TTA CCA CAC CTT	144
Pro Pro Gly Pro Thr Pro Trp Pro Ile Val Gly Asn Leu Pro His Leu	
35 40 45	
GGC CCA ATC CCA CAC CAC GCA CTC GCG GCC TTA GCC AAG AAG TAC GGG	192
Gly Pro Ile Pro His His Ala Leu Ala Ala Leu Ala Lys Lys Tyr Gly	
50 55 60	
CCA TTG ATG CAC CTG CGG CTC GGG TGT GTG GAC GTG GTT GTG GCC GCG	240
Pro Leu Met His Leu Arg Leu Gly Cys Val Asp Val Val Val Ala Ala	
65 70 75	
TCT GCT TCC GTA GCT GCA CAG TTT TTA AAA GTT CAC GAC GCA AAT TTT	288
Ser Ala Ser Val Ala Ala Gln Phe Leu Lys Val His Asp Ala Asn Phe	
80 85 90 95	
GCT AGT AGG CCG CCA AAT TCT GGC GCG AAA CAT GTG GCG TAT AAT TAT	336
Ala Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr	
100 105 110	
CAG GAT CTT GTG TTT GCA CCT TAT GGT CCA AGG TGG CGT TTG TTA AGG	384
Gln Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Leu Leu Arg	
115 120 125	
AAG ATT TGT TCG GTC CAT TTG TTT TCT GCT AAA GCA CTT GAT GAT TTT	432
Lys Ile Cys Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe	
130 135 140	
CGT CAT GTT CGA CAG GAG GAG GTA GCA GTC CTA ACC CGC GTA CTA CTG	480
Arg His Val Arg Gln Glu Glu Val Ala Val Leu Thr Arg Val Leu Leu	
145 150 155	
AGT GCT GGA AAC TCA CCG GTA CAG CTT GGC CAA CTA CTT AAC GTG TGT	528
Ser Ala Gly Asn Ser Pro Val Gln Leu Gly Gln Leu Leu Asn Val Cys	

- 158 -

160	165	170	175	
GCC ACA AAC GCC TTA GCA CGG GTA ATG TTA GGT AGG AGA GTT TTC GGA				576
Ala Thr Asn Ala Leu Ala Arg Val Met Leu Gly Arg Arg Val Phe Gly				
180	185	190		
GAC GGA ATT GAC AGG TCA GCC AAT GAG TTC AAA GAT ATG GTA GTA GAG				624
Asp Gly Ile Asp Arg Ser Ala Asn Glu Phe Lys Asp Met Val Val Glu				
195	200	205		
TTA ATG GTA TTA GCA GGA GAA TTT AAC CTT GGT GAC TTT ATT CCT GTA				672
Leu Met Val Leu Ala Gly Glu Phe Asn Leu Gly Asp Phe Ile Pro Val				
210	215	220		
CTT GAC CTA TTC GAC CTA CAA GGC ATT ACT AAA AAA ATG AAG AAG CTT				720
Leu Asp Leu Phe Asp Leu Gln Gly Ile Thr Lys Lys Met Lys Lys Leu				
225	230	235		
CAT GTT CGG TTC GAT TCA TTT CTT AGT AAG ATC GTT GAG GAG CAT AAA				768
His Val Arg Phe Asp Ser Phe Leu Ser Lys Ile Val Glu Glu His Lys				
240	245	250	255	
ACG GCA CCT GGT GGG TTG GGT CAT ACT GAT TTG CTG AGC ACG TTG ATT				816
Thr Ala Pro Gly Gly Leu Gly His Thr Asp Leu Leu Ser Thr Leu Ile				
260	265	270		
TCA CTT AAA GAT GAT GCT GAT ATT GAA GGT GGG AAG CTT ACA GAT ACT				864
Ser Leu Lys Asp Asp Ala Asp Ile Glu Gly Gly Lys Leu Thr Asp Thr				
275	280	285		
GAA ATC AAA GCT TTG CTT CTG AAT TTA TTC GCT GCG GGA ACA GAC ACA				912
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ala Ala Gly Thr Asp Thr				
290	295	300		
TCC TCT AGT ACA GTA GAA TGG GCA ATA GCC GAA CTC ATT CGT CAT CCA				960
Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg His Pro				
305	310	315		

160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315

- 159 -

[illegible]

- 160 -

Asp Trp Glu Leu Ala Asn Gly Leu Glu Pro Glu Met Leu Asn Met Glu
 465 470 475

GAA GCG TAT GGA TTG ACC CTT CAA CGG GCT GCA CCC TTG ATG GTT CAC 1488
 Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ala Pro Leu Met Val His
 480 485 490 495

CCG AAG CCG AGG TTA GCT CCC CAC GTA TAT GAA AGT ATT T AAGGACTAGT 1538
 Pro Lys Pro Arg Leu Ala Pro His Val Tyr Glu Ser Ile
 500 505

TTCTCTTTTG CCTTTTGTGTT TCGCAAAGGT TAATGAATAA ACGATTTCAT GACTCAGATA 1598

GTTATGTAAA CAATTGTGTT TGCTGTTTAT ATATTTATCT ATTTTCTAG AACAAAAAAA 1658

AA 1660

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 508 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Thr Ile Leu Ala Phe Val Phe Tyr Ala Leu Ile Leu Gly Ser Val
 1 5 10 15

Leu Tyr Val Phe Leu Asn Leu Ser Ser Arg Lys Ser Ala Arg Leu Pro
 20 25 30

Pro Gly Pro Thr Pro Trp Pro Ile Val Gly Asn Leu Pro His Leu Gly
 35 40 45

- 161 -

Pro Ile Pro His His Ala Leu Ala Ala Leu Ala Lys Lys Tyr Gly Pro
 50 55 60

Leu Met His Leu Arg Leu Gly Cys Val Asp Val Val Val Ala Ala Ser
 65 70 75 80

Ala Ser Val Ala Ala Gln Phe Leu Lys Val His Asp Ala Asn Phe Ala
 85 90 95

Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr Gln
 100 105 110

Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Leu Leu Arg Lys
 115 120 125

Ile Cys Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe Arg
 130 135 140

His Val Arg Gln Glu Glu Val Ala Val Leu Thr Arg Val Leu Leu Ser
 145 150 155 160

Ala Gly Asn Ser Pro Val Gln Leu Gly Gln Leu Leu Asn Val Cys Ala
 165 170 175

Thr Asn Ala Leu Ala Arg Val Met Leu Gly Arg Arg Val Phe Gly Asp
 180 185 190

Gly Ile Asp Arg Ser Ala Asn Glu Phe Lys Asp Met Val Val Glu Leu
 195 200 205

Met Val Leu Ala Gly Glu Phe Asn Leu Gly Asp Phe Ile Pro Val Leu
 210 215 220

Asp Leu Phe Asp Leu Gln Gly Ile Thr Lys Lys Met Lys Lys Leu His
 225 230 235 240

Val Arg Phe Asp Ser Phe Leu Ser Lys Ile Val Glu Glu His Lys Thr

- 162 -

245	250	255
Ala Pro Gly Gly Leu Gly His Thr Asp Leu Leu Ser Thr Leu Ile Ser		
260	265	270
Leu Lys Asp Asp Ala Asp Ile Glu Gly Gly Lys Leu Thr Asp Thr Glu		
275	280	285
Ile Lys Ala Leu Leu Leu Asn Leu Phe Ala Ala Gly Thr Asp Thr Ser		
290	295	300
Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg His Pro Gln		
305	310	315
Ile Leu Lys Gln Ala Arg Glu Glu Ile Asp Ala Val Val Gly Gln Asp		
325	330	335
Arg Leu Val Thr Glu Leu Asp Leu Ser Gln Leu Thr Tyr Leu Gln Ala		
340	345	350
Leu Val Lys Glu Val Phe Arg Leu His Pro Ser Thr Pro Leu Ser Leu		
355	360	365
Pro Arg Ile Ser Ser Glu Ser Cys Glu Val Asp Gly Tyr Tyr Ile Pro		
370	375	380
Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp Pro		
385	390	395
Lys Met Trp Ala Asp Pro Leu Glu Phe Arg Pro Ser Arg Phe Leu Pro		
405	410	415
Gly Gly Glu Lys Pro Gly Ala Asp Val Arg Gly Asn Asp Phe Glu Val		
420	425	430
Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu Gly		
435	440	445

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

- 163 -

Leu Arg Met Val Gln Leu Leu Ile Ala Thr Leu Val Gln Thr Phe Asp
 450 455 460

Trp Glu Leu Ala Asn Gly Leu Glu Pro Glu Met Leu Asn Met Glu Glu
 465 470 475 480

Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ala Pro Leu Met Val His Pro
 485 490 495

Lys Pro Arg Leu Ala Pro His Val Tyr Glu Ser Ile
 500 505

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1815 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 107..1631

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTAAATTAAT TAATAAATAC ACACACGACG AGATGTATGT AATGTAATGT AATATTATTA 60

CATACATCAT CACCGAATAC GCACGCTACT ACCACTGCGA TTAGCC ATG AGT CCC 115
 Met Ser Pro

1

TTA GCC TTG ATG ATC ATA AGT ACC TTA TTA GGG TTT CTC CTA TAC CAC 163

- 164 -

Leu Ala Leu Met Ile Ile Ser Thr Leu Leu Gly Phe Leu Leu Tyr His
 5 10 15
 TCT CTT CGC TTA CTA CTC TTC TCC GGC CAA GGT CGC CGA CTA CTA CCA 211
 Ser Leu Arg Leu Leu Leu Phe Ser Gly Gln Gly Arg Arg Leu Leu Pro
 20 25 30 35
 CCA GGT CCA CGC CCG TGG CCG CTG GTG GGA AAT CTC CCG CAC TTA GGC 259
 Pro Gly Pro Arg Pro Trp Pro Leu Val Gly Asn Leu Pro His Leu Gly
 40 45 50
 CCG AAG CCA CAC GCC TCC ATG GCC GAG CTC GCG CGA GCC TAC GGA CCC 307
 Pro Lys Pro His Ala Ser Met Ala Glu Leu Ala Arg Ala Tyr Gly Pro
 55 60 65
 CTC ATG CAC CTA AAG ATG GGG TTC GTC CAC GTC GTG GTG GCT TCG TCG 355
 Leu Met His Leu Lys Met Gly Phe Val His Val Val Val Ala Ser Ser
 70 75 80
 GCG AGC GCG GCG GAG CAG TGC CTG AGG GTT CAC GAC GCG AAT TTC TTG 403
 Ala Ser Ala Ala Glu Gln Cys Leu Arg Val His Asp Ala Asn Phe Leu
 85 90 95
 AGC AGG CCA CCC AAC TCC GGC GCC AAG CAC GTC GCT TAC AAC TAC GAG 451
 Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr Asn Tyr Glu
 100 105 110 115
 GAC TTG GTT TTC AGA CCG TAC GGT CCC AAG TGG AGG CTG TTG AGG AAG 499
 Asp Leu Val Phe Arg Pro Tyr Gly Pro Lys Trp Arg Leu Leu Arg Lys
 120 125 130
 ATA TGC GCT CAG CAT ATT TTC TCC GTC AAG GCT ATG GAT GAC TTC AGG 547
 Ile Cys Ala Gln His Ile Phe Ser Val Lys Ala Met Asp Asp Phe Arg
 135 140 145
 CGC GTC AGA GAG GAA GAG GTG GCC ATC CTG AGT CGC GCT CTA GCA GGC 595
 Arg Val Arg Glu Glu Glu Val Ala Ile Leu Ser Arg Ala Leu Ala Gly

- 165 -

150	155	160	
AAA AGG GCC GTA CCC ATA GGC CAA ATG CTC AAC GTG TGC GCC ACA AAC			643
Lys Arg Ala Val Pro Ile Gly Gln Met Leu Asn Val Cys Ala Thr Asn			
165	170	175	
GCC CTA TCT CGC GTC ATG ATG GGG CGG CGC GTG GTG GGC CAC GCG GAT			691
Ala Leu Ser Arg Val Met Met Gly Arg Arg Val Val Gly His Ala Asp			
180	185	190	195
GGA ACC AAC GAC GCC AAG GCG GAG GAG TTC AAA GCC ATG GTC GTC GAG			739
Gly Thr Asn Asp Ala Lys Ala Glu Glu Phe Lys Ala Met Val Val Glu			
200	205	210	
CTC ATG GTC CTC TCC GGC GTC TTC AAC ATC GGT GAT TTC ATC CCC TTC			787
Leu Met Val Leu Ser Gly Val Phe Asn Ile Gly Asp Phe Ile Pro Phe			
215	220	225	
CTC GAG CCT CTC GAC TTG CAG GGA GTG GCT TCC AAG ATG AAG AAA CTC			835
Leu Glu Pro Leu Asp Leu Gln Gly Val Ala Ser Lys Met Lys Lys Leu			
230	235	240	
CAC GCG CGG TTC GAT GCA TTC TTG ACC GAG ATT GTA CGA GAG CGT TGT			883
His Ala Arg Phe Asp Ala Phe Leu Thr Glu Ile Val Arg Glu Arg Cys			
245	250	255	
CAT GGG CAG ATC AAC AAC AGT GGT GCT CAT CAG GAT GAT TTG CTT AGC			931
His Gly Gln Ile Asn Asn Ser Gly Ala His Gln Asp Asp Leu Leu Ser			
260	265	270	275
ACG TTG ATT TCG TTC AAA GGG CTT GAC GAT GGC GAT GGT TCC AGG CTC			979
Thr Leu Ile Ser Phe Lys Gly Leu Asp Asp Gly Asp Gly Ser Arg Leu			
280	285	290	
ACT GAC ACA GAA ATC AAG GCG CTG CTC TTG AAC CTT TTG GAC ACG ACG			1027
Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Leu Leu Asp Thr Thr			
295	300	305	

- 166 -

TCG AGC ACG GTG GAA TGG GCC GTA GCC GAA CTC CTA CGC CAC CCT AAG	1075
Ser Ser Thr Val Glu Trp Ala Val Ala Glu Leu Leu Arg His Pro Lys	
310 315 320	
ACA TTA GCC CAA GTC CGG CAA GAG CTC GAC TCG GTC GTG GGT AAG AAC	1123
Thr Leu Ala Gln Val Arg Gln Glu Leu Asp Ser Val Val Gly Lys Asn	
325 330 335	
AGG CTC GTG TCC GAG ACC GAT CTG AAT CAG CTG CCC TAT CTA CAA GCT	1171
Arg Leu Val Ser Glu Thr Asp Leu Asn Gln Leu Pro Tyr Leu Gln Ala	
340 345 350 355	
GTC GTC AAA GAA ACT TTC CGC CTC CAT CCT CCG ACG CCG CTC TCT CTA	1219
Val Val Lys Glu Thr Phe Arg Leu His Pro Pro Thr Pro Leu Ser Leu	
360 365 370	
CCG AGA CTC GCG GAA GAT GAT TGC GAG ATC GAC GGA TAC CTC ATC CCC	1267
Pro Arg Leu Ala Glu Asp Asp Cys Glu Ile Asp Gly Tyr Leu Ile Pro	
375 380 385	
AAG GGC TCG ACC CTT CTG GTG AAC GTT TGG GCC ATA GCC CGC GAT CCC	1315
Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp Pro	
390 395 400	
AAG GTT TGG GCC GAT CCG TTG GAG TTT AGG CCC GAA CGA TTC TTG ACG	1363
Lys Val Trp Ala Asp Pro Leu Glu Phe Arg Pro Glu Arg Phe Leu Thr	
405 410 415	
GGC GGA GAA AAG GCC GAC GTC GAT GTC AAG GGG AAC GAT TTC GAA GTG	1411
Gly Gly Glu Lys Ala Asp Val Asp Val Lys Gly Asn Asp Phe Glu Val	
420 425 430 435	
ATA CCG TTC GGG GCG GGT CGT AGG ATC TGC GCT GGC GTT GGC TTG GGA	1459
Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Val Gly Leu Gly	
440 445 450	
ATA CGT ATG GTC CAA CTG TTG ACG GCG AGT TTG ATC CAT GCA TTC GAT	1507

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

- 167 -

Ile Arg Met Val Gln Leu Leu Thr Ala Ser Leu Ile His Ala Phe Asp

455

460

465

CTG GAC CTT GCT AAT GGG CTT TTG GCC CAA AAT CTG AAC ATG GAA GAA 1555

Leu Asp Leu Ala Asn Gly Leu Leu Ala Gln Asn Leu Asn Met Glu Glu

470

475

480

GCA TAT GGG CTT ACG CTA CAA CGG GCT GAG CCT TTG TTG GTC CAC CCT 1603

Ala Tyr Gly Leu Thr Leu Gln Arg Ala Glu Pro Leu Leu Val His Pro

485

490

495

AGG CCG CGG TTG GCC ACT CAT GTC TAT T AATTAAATTA GGCCTAAACT 1651

Arg Pro Arg Leu Ala Thr His Val Tyr

500

505

ACGATGAATG ACCCATTTAA CGTTAATAAG AGTTTTCAAT TTATGTGAGT TTGCATGGTA 1711

TGGTATGGTA TGGTGCTTGT AATAAATTGT ATCTGTTAGG TGTGTTTCATT GATGATAAAT 1771

CTAGTTTGTA CTGCTGCTCA AAAAAAAAAA AAAAAAAAAA AAAA 1815

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 508 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Ser Pro Leu Ala Leu Met Ile Ile Ser Thr Leu Leu Gly Phe Leu

1

5

10

15

Leu Tyr His Ser Leu Arg Leu Leu Leu Phe Ser Gly Gln Gly Arg Arg

- 168 -

	20	25	30
Leu Leu Pro Pro Gly Pro Arg Pro Trp Pro Leu Val Gly Asn Leu Pro	35	40	45
His Leu Gly Pro Lys Pro His Ala Ser Met Ala Glu Leu Ala Arg Ala	50	55	60
Tyr Gly Pro Leu Met His Leu Lys Met Gly Phe Val His Val Val Val	65	70	75 80
Ala Ser Ser Ala Ser Ala Ala Glu Gln Cys Leu Arg Val His Asp Ala	85	90	95
Asn Phe Leu Ser Arg Pro Pro Asn Ser Gly Ala Lys His Val Ala Tyr	100	105	110
Asn Tyr Glu Asp Leu Val Phe Arg Pro Tyr Gly Pro Lys Trp Arg Leu	115	120	125
Leu Arg Lys Ile Cys Ala Gln His Ile Phe Ser Val Lys Ala Met Asp	130	135	140
Asp Phe Arg Arg Val Arg Glu Glu Glu Val Ala Ile Leu Ser Arg Ala	145	150	155 160
Leu Ala Gly Lys Arg Ala Val Pro Ile Gly Gln Met Leu Asn Val Cys	165	170	175
Ala Thr Asn Ala Leu Ser Arg Val Met Met Gly Arg Arg Val Val Gly	180	185	190
His Ala Asp Gly Thr Asn Asp Ala Lys Ala Glu Glu Phe Lys Ala Met	195	200	205
Val Val Glu Leu Met Val Leu Ser Gly Val Phe Asn Ile Gly Asp Phe	210	215	220

- 169 -

Ile Pro Phe Leu Glu Pro Leu Asp Leu Gln Gly Val Ala Ser Lys Met
225 230 235 240

Lys Lys Leu His Ala Arg Phe Asp Ala Phe Leu Thr Glu Ile Val Arg
245 250 255

Glu Arg Cys His Gly Gln Ile Asn Asn Ser Gly Ala His Gln Asp Asp
260 265 270

Leu Leu Ser Thr Leu Ile Ser Phe Lys Gly Leu Asp Asp Gly Asp Gly
275 280 285

Ser Arg Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn Leu Leu
290 295 300

Asp Thr Thr Ser Ser Thr Val Glu Trp Ala Val Ala Glu Leu Leu Arg
305 310 315 320

His Pro Lys Thr Leu Ala Gln Val Arg Gln Glu Leu Asp Ser Val Val
325 330 335

Gly Lys Asn Arg Leu Val Ser Glu Thr Asp Leu Asn Gln Leu Pro Tyr
340 345 350

Leu Gln Ala Val Val Lys Glu Thr Phe Arg Leu His Pro Pro Thr Pro
355 360 365

Leu Ser Leu Pro Arg Leu Ala Glu Asp Asp Cys Glu Ile Asp Gly Tyr
370 375 380

Leu Ile Pro Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala
385 390 395 400

Arg Asp Pro Lys Val Trp Ala Asp Pro Leu Glu Phe Arg Pro Glu Arg
405 410 415

Phe Leu Thr Gly Gly Glu Lys Ala Asp Val Asp Val Lys Gly Asn Asp

- 170 -

420	425	430
Phe Glu Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Val		
435	440	445
Gly Leu Gly Ile Arg Met Val Gln Leu Leu Thr Ala Ser Leu Ile His		
450	455	460
Ala Phe Asp Leu Asp Leu Ala Asn Gly Leu Leu Ala Gln Asn Leu Asn		
465	470	475
Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Glu Pro Leu Leu		
485	490	495
Val His Pro Arg Pro Arg Leu Ala Thr His Val Tyr		
500	505	

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1824 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..1553

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

G AGC TTA ACC TTA ATT TTC TGC ACT TTA GTT TTT GCA ATC TTT CTA
 Ser Leu Thr Leu Ile Phe Cys Thr Leu Val Phe Ala Ile Phe Leu

- 171 -

1	5	10	15	
TAT TTT CTT ATT CTC AGG GTG AAA CAG CGT TAC CCT TTA CCT CTC CCA				94
Tyr Phe Leu Ile Leu Arg Val Lys Gln Arg Tyr Pro Leu Pro Leu Pro				
	20	25	30	
CCC GGA CCA AAA CCA TGG CCG GTG TTA GGA AAC CTT CCC CAC CTG GGC				142
Pro Gly Pro Lys Pro Trp Pro Val Leu Gly Asn Leu Pro His Leu Gly				
	35	40	45	
AAG AAG CCT CAC CAG TCG ATT GCG GCC ATG GCT GAG AGG TAC GGC CCC				190
Lys Lys Pro His Gln Ser Ile Ala Ala Met Ala Glu Arg Tyr Gly Pro				
	50	55	60	
CTC ATG CAC CTC CGC CTA GGA TTC GTG GAC GTG GTT GTG GCC GCC TCC				238
Leu Met His Leu Arg Leu Gly Phe Val Asp Val Val Val Ala Ala Ser				
	65	70	75	
GCC GCC GTG GCC GCT CAG TTC TTG AAA GTT CAC GAC TCG AAC TTC TCC				286
Ala Ala Val Ala Ala Gln Phe Leu Lys Val His Asp Ser Asn Phe Ser				
	80	85	90	95
AAC CGG CCG CCG AAC TCC GGC GCG GAA CAC ATT GCT TAT AAC TAT CAA				334
Asn Arg Pro Pro Asn Ser Gly Ala Glu His Ile Ala Tyr Asn Tyr Gln				
	100	105	110	
GAC CTC GTC TTC GCG CCC TAC GGC CCG CGG TGG CGC ATG CTT AGG AAG				382
Asp Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu Arg Lys				
	115	120	125	
ATC ACC TCC GTG CAT CTC TTC TCG GCC AAG GCG TTG GAT GAC TTC TGC				430
Ile Thr Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe Cys				
	130	135	140	
CAT GTT CGC CAG GAA GAG GTT GCA ACT CTG ACA CGC AGT CTA GCA AGT				478
His Val Arg Gln Glu Glu Val Ala Thr Leu Thr Arg Ser Leu Ala Ser				
	145	150	155	

- 172 -

GCA GGC AAA ACT CCA GTA AAA CTA GGG CAG TTA CTA AAC GTG TGC ACC 526
Ala Gly Lys Thr Pro Val Lys Leu Gly Gln Leu Leu Asn Val Cys Thr
160 165 170 175

ACG AAC GCC CTA GCT CGT GTA ATG CTA GGG CGG AAG GTC TTT AAT GAC 574
Thr Asn Ala Leu Ala Arg Val Met Leu Gly Arg Lys Val Phe Asn Asp
180 185 190

GGA GGT AGC AAG AGC GAC CCA AAG GCG GAG GAG TTC AAG TCG ATG GTG 622
Gly Gly Ser Lys Ser Asp Pro Lys Ala Glu Glu Phe Lys Ser Met Val
195 200 205

GAG GAG ATG ATG GTG TTG GCC GGA AGT TTT AAC ATC GGC GAT TTC ATT 670
Glu Glu Met Met Val Leu Ala Gly Ser Phe Asn Ile Gly Asp Phe Ile
210 215 220

CCG GTC TTG GGT TGG TTT GAC GTT CAG GGT ATC GTA GGG AAG ATG AAG 718
Pro Val Leu Gly Trp Phe Asp Val Gln Gly Ile Val Gly Lys Met Lys
225 230 235

AAA CTA CAC GCG CGT TTT GAT GCG TTC TTG AAC ACC ATT CTA GAG GAA 766
Lys Leu His Ala Arg Phe Asp Ala Phe Leu Asn Thr Ile Leu Glu Glu
240 245 250 255

CAC AAA TGT GTC AAC AAT CAA CAC ACG ACG TTG TCG AAA GAT GTG GAC 814
His Lys Cys Val Asn Asn Gln His Thr Thr Leu Ser Lys Asp Val Asp
260 265 270

TTC TTG AGC ACC CTA ATT AGG CTC AAA GAT AAT GGG GCT GAT ATG GAT 862
Phe Leu Ser Thr Leu Ile Arg Leu Lys Asp Asn Gly Ala Asp Met Asp
275 280 285

TGT GAA GAG GGA AAA CTC ACC GAC ACT GAA ATT AAG GCT TTG CTC TTG 910
Cys Glu Glu Gly Lys Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu
290 295 300

AAC CTG TTC ACA GCT GGG ACT GAT ACA TCA TCT AGC ACT GTG GAG TGG 958

Asn Leu Phe Thr Ala Gly Thr Asp Thr Ser Ser Ser Thr Val Glu Trp
305 310 315

CAA GAG CTT GAC TTA GTG GTG GGT CAA AAT CAG CTA GTC ACA GAA TCT 1054
Gln Glu Leu Asp Leu Val Val Gly Gln Asn Gln Leu Val Thr Glu Ser
340 345 350

AGG CTA CAC CCA TCC ACC CCA CTC TCT CTT CCA AGA ATG GGA GCT CAG 1150
Arg Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Met Gly Ala Gln
370 375 380

GTC AAC GTT TGG GCC ATA GCT CGT GAT CCC AAT GTG TGG ACA AAT CCT 1246
Val Asn Val Trp Ala Ile Ala Arg Asp Pro Asn Val Trp Thr Asn Pro
400 405 410 415

GTG	GAT	ATT	AAA	GGG	AAT	GAC	TTT	GAA	GTG	ATT	CCT	TTT	GGA	GCC	GGG	1342
Val	Asp	Ile	Lys	Gly	Asn	Asp	Phe	Glu	Val	Ile	Pro	Phe	Gly	Ala	Gly	
			435				440						445			

CGT AGA ATA TGC TCT GGG ATG AGT TTG GGG ATA AGG ATG GTT CAC CTG 1390
Arg Arg Ile Cys Ser Gly Met Ser Leu Gly Ile Arg Met Val His Leu

- 174 -

450	455	460	
TTG GTT GCA ACT TTG GTG CAT GCT TTT GAT TGG GAT TTG GTG AAT GGA			1438
Leu Val Ala Thr Leu Val His Ala Phe Asp Trp Asp Leu Val Asn Gly			
465	470	475	
CAA TCT GTA GAG ACG CTC AAT ATG GAG GAA GCT TAT GGT CTC ACC CTT			1486
Gln Ser Val Glu Thr Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu			
480	485	490	495
CAA CGA GCT GTT CCT TTG ATG TTG CAT CCA AAG CCC AGA TTA CAA CCA			1534
Gln Arg Ala Val Pro Leu Met Leu His Pro Lys Pro Arg Leu Gln Pro			
	500	505	510
CAT CTC TAT ACT CTC AAT T AAATTGCAAT TTGATTTTGG TGATTATACA			1583
His Leu Tyr Thr Leu Asn			
	515		
ATTATAATCG AGGGACATAG GATCCCCATT TATTTATATT CAGTTATAAG AGACTTCCAA			1643
CAAAGGTCTA GCTTTCGACC TTAAAAGTTG TAAAAGAGGT CCTACATATG TAAAAGCCCG			1703
CCAAAGGAAA ACTGGTTGTA TTCAATTCCG CTAGGCCTTG TCCGAAAGAC CTCATGAAGA			1763
CTACAAAGGT CATATATAAT GGTAACCCA GTGTATTTGT TGTAAAAAAA AAAAAAAAAA			1823
A			1824

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 517 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 175 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ser Leu Thr Leu Ile Phe Cys Thr Leu Val Phe Ala Ile Phe Leu Tyr
 1 5 10 15

Phe Leu Ile Leu Arg Val Lys Gln Arg Tyr Pro Leu Pro Leu Pro Pro
 20 25 30

Gly Pro Lys Pro Trp Pro Val Leu Gly Asn Leu Pro His Leu Gly Lys
 35 40 45

Lys Pro His Gln Ser Ile Ala Ala Met Ala Glu Arg Tyr Gly Pro Leu
 50 55 60

Met His Leu Arg Leu Gly Phe Val Asp Val Val Val Ala Ala Ser Ala
 65 70 75 80

Ala Val Ala Ala Gln Phe Leu Lys Val His Asp Ser Asn Phe Ser Asn
 85 90 95

Arg Pro Pro Asn Ser Gly Ala Glu His Ile Ala Tyr Asn Tyr Gln Asp
 100 105 110

Leu Val Phe Ala Pro Tyr Gly Pro Arg Trp Arg Met Leu Arg Lys Ile
 115 120 125

Thr Ser Val His Leu Phe Ser Ala Lys Ala Leu Asp Asp Phe Cys His
 130 135 140

Val Arg Gln Glu Glu Val Ala Thr Leu Thr Arg Ser Leu Ala Ser Ala
 145 150 155 160

Gly Lys Thr Pro Val Lys Leu Gly Gln Leu Leu Asn Val Cys Thr Thr
 165 170 175

Asn Ala Leu Ala Arg Val Met Leu Gly Arg Lys Val Phe Asn Asp Gly
 180 185 190

- 176 -

Gly Ser Lys Ser Asp Pro Lys Ala Glu Glu Phe Lys Ser Met Val Glu
 195 200 205

Glu Met Met Val Leu Ala Gly Ser Phe Asn Ile Gly Asp Phe Ile Pro
 210 215 220

Val Leu Gly Trp Phe Asp Val Gln Gly Ile Val Gly Lys Met Lys Lys
 225 230 235 240

Leu His Ala Arg Phe Asp Ala Phe Leu Asn Thr Ile Leu Glu Glu His
 245 250 255

Lys Cys Val Asn Asn Gln His Thr Thr Leu Ser Lys Asp Val Asp Phe
 260 265 270

Leu Ser Thr Leu Ile Arg Leu Lys Asp Asn Gly Ala Asp Met Asp Cys
 275 280 285

Glu Glu Gly Lys Leu Thr Asp Thr Glu Ile Lys Ala Leu Leu Leu Asn
 290 295 300

Leu Phe Thr Ala Gly Thr Asp Thr Ser Ser Ser Thr Val Glu Trp Ala
 305 310 315 320

Ile Ala Glu Leu Leu Arg Asn Pro Lys Ile Leu Asn Gln Ala Gln Gln
 325 330 335

Glu Leu Asp Leu Val Val Gly Gln Asn Gln Leu Val Thr Glu Ser Asp
 340 345 350

Leu Thr Asp Leu Pro Phe Leu Gln Ala Ile Val Lys Glu Thr Phe Arg
 355 360 365

Leu His Pro Ser Thr Pro Leu Ser Leu Pro Arg Met Gly Ala Gln Gly
 370 375 380

Cys Glu Ile Asn Gly Tyr Phe Ile Pro Lys Gly Ala Thr Leu Leu Val

- 177 -

385 390 395 400
 Asn Val Trp Ala Ile Ala Arg Asp Pro Asn Val Trp Thr Asn Pro Leu
 405 410 415
 Glu Phe Asn Pro His Arg Phe Leu Pro Gly Gly Glu Lys Pro Asn Val
 420 425 430
 Asp Ile Lys Gly Asn Asp Phe Glu Val Ile Pro Phe Gly Ala Gly Arg
 435 440 445
 Arg Ile Cys Ser Gly Met Ser Leu Gly Ile Arg Met Val His Leu Leu
 450 455 460
 Val Ala Thr Leu Val His Ala Phe Asp Trp Asp Leu Val Asn Gly Gln
 465 470 475 480
 Ser Val Glu Thr Leu Asn Met Glu Glu Ala Tyr Gly Leu Thr Leu Gln
 485 490 495
 Arg Ala Val Pro Leu Met Leu His Pro Lys Pro Arg Leu Gln Pro His
 500 505 510
 Leu Tyr Thr Leu Asn
 515

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1667 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 178 -

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1429

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCC ATC CTC GGA AAC ATC CCC CAT CTC GGC TCC AAA CCG CAC CAA ACA	48
Pro Ile Leu Gly Asn Ile Pro His Leu Gly Ser Lys Pro His Gln Thr	
1 5 10 15	
CTC GCG GAA ATG GCG AAA ACC TAC GGT CCG CTC ATG CAC TTG AAG TTC	96
Leu Ala Glu Met Ala Lys Thr Tyr Gly Pro Leu Met His Leu Lys Phe	
20 25 30	
GGG CTT AAG GAC GCG GTG GTG GCG TCG TCT GCG TCG GTG GCA GAG CAG	144
Gly Leu Lys Asp Ala Val Val Ala Ser Ser Ala Ser Val Ala Glu Gln	
35 40 45	
TTT CTG AAG AAA CAC GAC GTG AAT TTC TCG AAC CGG CCG CCA AAC TCC	192
Phe Leu Lys Lys His Asp Val Asn Phe Ser Asn Arg Pro Pro Asn Ser	
50 55 60	
GGG GCC AAA CAT ATA GCT TAT AAC TAT CAG GAC CTG GTA TTC GCT CCC	240
Gly Ala Lys His Ile Ala Tyr Asn Tyr Gln Asp Leu Val Phe Ala Pro	
65 70 75 80	
TAT GGA CCC CGG TGG CGG TTG CTT AGG AAA ATC TGT TCC GTC CAT CTT	288
Tyr Gly Pro Arg Trp Arg Leu Leu Arg Lys Ile Cys Ser Val His Leu	
85 90 95	
TTC TCG TCT AAG GCC TTG GAT GAC TTT CAG CAT GTT CGA CAT GAG GAG	336
Phe Ser Ser Lys Ala Leu Asp Asp Phe Gln His Val Arg His Glu Glu	
100 105 110	
ATA TGC ATC CTT ATA CGA GCA ATA GCG AGT GGC GGT CAT GCT CCG GTG	384
Ile Cys Ile Leu Ile Arg Ala Ile Ala Ser Gly Gly His Ala Pro Val	

- 179 -

115	120	125	
AAT TTA GGC AAG TTA TTA GGA GTG TGC ACA ACC AAT GCC CTG GCA AGA			432
Asn Leu Gly Lys Leu Leu Gly Val Cys Thr Thr Asn Ala Leu Ala Arg			
130	135	140	
GTG ATG CTT GGA AGA AGA GTA TTC GAA GGC GAC GGC GGC GAG AAT CCG			480
Val Met Leu Gly Arg Arg Val Phe Glu Gly Asp Gly Gly Glu Asn Pro			
145	150	155	160
CAT GCC GAC GAG TTT AAA TCA ATG GTG GTG GAG ATT ATG GTG TTA GCC			528
His Ala Asp Glu Phe Lys Ser Met Val Val Glu Ile Met Val Leu Ala			
165	170	175	
GGT GCA TTC AAC TTG GGT GAT TTC ATC CCG GTT CTA GAT TGG TTC GAT			576
Gly Ala Phe Asn Leu Gly Asp Phe Ile Pro Val Leu Asp Trp Phe Asp			
180	185	190	
TTG CAA GGA ATT GCT GGT AAA ATG AAG AAA CTT CAT GCC CGT TTC GAC			624
Leu Gln Gly Ile Ala Gly Lys Met Lys Lys Leu His Ala Arg Phe Asp			
195	200	205	
AAG TTT TTA AAT GGG ATC CTA GAA GAT CGT AAA TCT AAC GGC TCT AAT			672
Lys Phe Leu Asn Gly Ile Leu Glu Asp Arg Lys Ser Asn Gly Ser Asn			
210	215	220	
GGA GCT GAA CAA TAC GTG GAC TTG CTC AGT GTG TTG ATC TCT CTT CAA			720
Gly Ala Glu Gln Tyr Val Asp Leu Leu Ser Val Leu Ile Ser Leu Gln			
225	230	235	240
GAT AGT AAT ATC GAC GGT GGT GAC GAA GGA ACC AAA CTC ACA GAT ACT			768
Asp Ser Asn Ile Asp Gly Gly Asp Glu Gly Thr Lys Leu Thr Asp Thr			
245	250	255	
GAA ATC AAA GCT CTC CTT TTG AAC TTG TTC ATA GCC GGA ACA GAC ACT			816
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ile Ala Gly Thr Asp Thr			
260	265	270	

- 180 -

TCA TCA AGT ACT GTA GAA TGG GCC ATG GCA GAA CTA ATC CGA AAC CCA	864
Ser Ser Ser Thr Val Glu Trp Ala Met Ala Glu Leu Ile Arg Asn Pro	
275 280 285	
AAG TTA CTA GTC CAA GCC CAA GAA GAG CTA GAC AGA GTA GTC GGG CCG	912
Lys Leu Leu Val Gln Ala Gln Glu Glu Leu Asp Arg Val Val Gly Pro	
290 295 300	
AAC CGA TTC GTA ACC GAA TCT GAT CTT CCT CAA CTG ACA TTC CTT CAA	960
Asn Arg Phe Val Thr Glu Ser Asp Leu Pro Gln Leu Thr Phe Leu Gln	
305 310 315 320	
GCC GTC ATC AAA GAG ACT TTC AGG CTT CAT CCA TCC ACC CCA CTC TCT	1008
Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro Leu Ser	
325 330 335	
CTT CCA CGA ATG GCG GCG GAG GAC TGT GAG ATC AAT GGG TAT TAT GTC	1056
Leu Pro Arg Met Ala Ala Glu Asp Cys Glu Ile Asn Gly Tyr Tyr Val	
340 345 350	
TCA GAA GGT TCG ACA TTG CTC GTC AAT GTG TGG GCC ATA GCT CGT GAT	1104
Ser Glu Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp	
355 360 365	
CCA AAT GCG TGG GCC AAT CCA CTA GAT TTC AAC CCG ACT CGT TTC TTG	1152
Pro Asn Ala Trp Ala Asn Pro Leu Asp Phe Asn Pro Thr Arg Phe Leu	
370 375 380	
GCC GGT GGA GAG AAG CCT AAT GTT GAT GTT AAA GGG AAT GAT TTT GAA	1200
Ala Gly Gly Glu Lys Pro Asn Val Asp Val Lys Gly Asn Asp Phe Glu	
385 390 395 400	
GTG ATA CCT TTC GGT GCT GGG CGC AGG ATA TGT GCC GGA ATG AGC TTA	1248
Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu	
405 410 415	
GGT ATA CGG ATG GTT CAA CTA GTA ACG GCT TCG TTA GTT CAT TCG TTT	1296

- 181 -

Gly Ile Arg Met Val Gln Leu Val Thr Ala Ser Leu Val His Ser Phe	
420 425 430	
GAT TGG GCT TTG TTG GAT GGA CTT AAA CCC GAG AAG CTT GAC ATG GAG	1344
Asp Trp Ala Leu Leu Asp Gly Leu Lys Pro Glu Lys Leu Asp Met Glu	
435 440 445	
GAA GGT TAT GGA CTA ACG CTT CAA CGA GCT TCA CCT TTA ATC GTC CAT	1392
Glu Gly Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His	
450 455 460	
CCA AAG CCG AGG CTC TCG GCT CAA GTT TAT TGT ATG T AACAAGTTTG	1439
Pro Lys Pro Arg Leu Ser Ala Gln Val Tyr Cys Met	
465 470 475	
TGAAGCCAGT CTGATTTTCAG TTGGATTTGT AGTTATTTTA TGATCATTTG GTATTTTATT	1499
TTGTATTTTCG GTTGAATACA ATAAAGGGAA GGTGGATCGT CTGCTGTATA ATAGCGACGT	1559
TTTAACGTGT TGTGATAGTA CCGTGTTTTA CTAAACGAT GTCGTTTGAT TTTTATAGT	1619
ATTAAAAAAA TAAACAGCTG GATTTTGAAC CAAAAAAA AAAAAAAA	1667

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 476 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Pro Ile Leu Gly Asn Ile Pro His Leu Gly Ser Lys Pro His Gln Thr
1 5 10 15

- 183 -

210	215	220
Gly Ala Glu Gln Tyr Val Asp Leu Leu Ser Val Leu Ile Ser Leu Gln		
225	230	235 240
Asp Ser Asn Ile Asp Gly Gly Asp Glu Gly Thr Lys Leu Thr Asp Thr		
245	250	255
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Ile Ala Gly Thr Asp Thr		
260	265	270
Ser Ser Ser Thr Val Glu Trp Ala Met Ala Glu Leu Ile Arg Asn Pro		
275	280	285
Lys Leu Leu Val Gln Ala Gln Glu Glu Leu Asp Arg Val Val Gly Pro		
290	295	300
Asn Arg Phe Val Thr Glu Ser Asp Leu Pro Gln Leu Thr Phe Leu Gln		
305	310	315 320
Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro Leu Ser		
325	330	335
Leu Pro Arg Met Ala Ala Glu Asp Cys Glu Ile Asn Gly Tyr Tyr Val		
340	345	350
Ser Glu Gly Ser Thr Leu Leu Val Asn Val Trp Ala Ile Ala Arg Asp		
355	360	365
Pro Asn Ala Trp Ala Asn Pro Leu Asp Phe Asn Pro Thr Arg Phe Leu		
370	375	380
Ala Gly Gly Glu Lys Pro Asn Val Asp Val Lys Gly Asn Asp Phe Glu		
385	390	395 400
Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu		
405	410	415

- 184 -

Gly Ile Arg Met Val Gln Leu Val Thr Ala Ser Leu Val His Ser Phe
 420 425 430

Asp Trp Ala Leu Leu Asp Gly Leu Lys Pro Glu Lys Leu Asp Met Glu
 435 440 445

Glu Gly Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His
 450 455 460

Pro Lys Pro Arg Leu Ser Ala Gln Val Tyr Cys Met
 465 470 475

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1214 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..1091

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

T CGC ATC CTC ACG CGA TCT ATA GCG AGT GCT GGG GAA AAT CCG ATT	46
Arg Ile Leu Thr Arg Ser Ile Ala Ser Ala Gly Glu Asn Pro Ile	
1 5 10 15	
AAC TTA GGT CAA TTA CTC GGG GTG TGT ACC ACA AAT GCT CTG GCG AGA	94
Asn Leu Gly Gln Leu Leu Gly Val Cys Thr Thr Asn Ala Leu Ala Arg	
20 25 30	

- 185 -

GTG ATG CTT GGA AGG AGG GTA TTC GGC GAT GGG AGC GGC GGC GTA GAT	142
Val Met Leu Gly Arg Arg Val Phe Gly Asp Gly Ser Gly Gly Val Asp	
35 40 45	
CCT CAG GCG GAC GAG TTC AAA TCC ATG GTG GTG GAA ATC ATG GTG TTG	190
Pro Gln Ala Asp Glu Phe Lys Ser Met Val Val Glu Ile Met Val Leu	
50 55 60	
GCC GGC GCG TTT AAT CTA GGT GAT TTT ATT CCC GCT CTT GAT TGG TTC	238
Ala Gly Ala Phe Asn Leu Gly Asp Phe Ile Pro Ala Leu Asp Trp Phe	
65 70 75	
GAT CTG CAG GGA ATT ACG GCA AAA ATG AAG AAA GTT CAC GCT CGT TTC	286
Asp Leu Gln Gly Ile Thr Ala Lys Met Lys Lys Val His Ala Arg Phe	
80 85 90 95	
GAT GCG TTC TTA GAC GCG ATC CTT GAG GAG CAC AAA TCC AAC GGC TCT	334
Asp Ala Phe Leu Asp Ala Ile Leu Glu Glu His Lys Ser Asn Gly Ser	
100 105 110	
CGC GGA GCT AAG CAA CAC GTT GAC TTG CTG AGT ATG TTG ATC TCC CTT	382
Arg Gly Ala Lys Gln His Val Asp Leu Leu Ser Met Leu Ile Ser Leu	
115 120 125	
CAA GAT AAT AAC ATT GAT GGT GAA AGT GGC GCC AAA CTC ACT GAT ACA	430
Gln Asp Asn Asn Ile Asp Gly Glu Ser Gly Ala Lys Leu Thr Asp Thr	
130 135 140	
GAA ATC AAA GCT TTG CTT CTG AAC TTG TTC ACG GCT GGA ACA GAC ACG	478
Glu Ile Lys Ala Leu Leu Leu Asn Leu Phe Thr Ala Gly Thr Asp Thr	
145 150 155	
TCA TCA AGT ACT GTG GAG TGG GCA ATC GCA GAG CTA ATC CGA AAC CCA	526
Ser Ser Ser Thr Val Glu Trp Ala Ile Ala Glu Leu Ile Arg Asn Pro	
160 165 170 175	
GAA GTA TTG GTT CAA GCC CAA CAA GAG CTC GAT AGA GTA GTT GGG CCA	574

- 186 -

Glu Val Leu Val Gln Ala Gln Gln Glu Leu Asp Arg Val Val Gly Pro	
180 185 190	
AGT CGT CTT GTG ACC GAA TCT GAT CTG CCT CAA TTG GCA TTC CTT CAA	622
Ser Arg Leu Val Thr Glu Ser Asp Leu Pro Gln Leu Ala Phe Leu Gln	
195 200 205	
GCT GTC ATC AAA GAG ACT TTC AGA CTT CAT CCA TCC ACT CCA CTC TCT	670
Ala Val Ile Lys Glu Thr Phe Arg Leu His Pro Ser Thr Pro Leu Ser	
210 215 220	
CTT CCA CGA ATG GCT TCA GAG GGT TGT GAA ATC AAT GGA TAC TCC ATC	718
Leu Pro Arg Met Ala Ser Glu Gly Cys Glu Ile Asn Gly Tyr Ser Ile	
225 230 235	
CCA AAG GGT TCG ACA TTG CTC GTT AAC GTA TGG TCC ATA GCC CGT GAT	766
Pro Lys Gly Ser Thr Leu Leu Val Asn Val Trp Ser Ile Ala Arg Asp	
240 245 250 255	
CCT AGT ATA TGG GCC GAC CCA TTA GAA TTT AGG CCG GCA CGT TTC TTG	814
Pro Ser Ile Trp Ala Asp Pro Leu Glu Phe Arg Pro Ala Arg Phe Leu	
260 265 270	
CCC GGC GGA GAA AAG CCC AAT GTT GAT GTG AGA GGC AAT GAT TTT GAG	862
Pro Gly Gly Glu Lys Pro Asn Val Asp Val Arg Gly Asn Asp Phe Glu	
275 280 285	
GTC ATA CCA TTT GGT GCT GGA CGT AGG ATA TGT GCT GGA ATG AGC TTG	910
Val Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu	
290 295 300	
GGT TTA AGA ATG GTT CAA CTT TCG ACA GCT ACT TTG GTT CAT TCG TTT	958
Gly Leu Arg Met Val Gln Leu Ser Thr Ala Thr Leu Val His Ser Phe	
305 310 315	
AAT TGG GAT TTG CTG AAT GGG ATG AGC CCA GAT AAA CTT GAC ATG GAA	1006
Asn Trp Asp Leu Leu Asn Gly Met Ser Pro Asp Lys Leu Asp Met Glu	

- 187 -

320	325	330	335	
GAA GCT TAT GGG CTT ACA TTG CAA CGG GCT TCA CCT TTG ATT GTC CAC				1054
Glu Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His				
	340	345	350	
CCA AAG CCC AGG CTT GCT AGC TCT ATG TAT GTT AAA T GAAATTATGC				1101
Pro Lys Pro Arg Leu Ala Ser Ser Met Tyr Val Lys				
	355	360		
TGTGCGAATA ATTCCTTATT TATAGCAGGA AATGTCATCT TGAATTATGT GTAATGTTCT				1161
TCTAACTTTC GATGGAAGTG CAAAACAAGT TTTATTAAAA AAAAAAAAAA AAA				1214

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 363 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Arg	Ile	Leu	Thr	Arg	Ser	Ile	Ala	Ser	Ala	Gly	Glu	Asn	Pro	Ile	Asn
1					5					10					15
Leu	Gly	Gln	Leu	Leu	Gly	Val	Cys	Thr	Thr	Asn	Ala	Leu	Ala	Arg	Val
					20					25					30
Met	Leu	Gly	Arg	Arg	Val	Phe	Gly	Asp	Gly	Ser	Gly	Gly	Val	Asp	Pro
					35					40					45
Gln	Ala	Asp	Glu	Phe	Lys	Ser	Met	Val	Val	Glu	Ile	Met	Val	Leu	Ala
					50					55					60

[illegible]

Ser Ile Trp Ala Asp Pro Leu Glu Phe Arg Pro Ala Arg Phe Leu Pro

- 189 -

260	265	270
Gly Gly Glu Lys Pro Asn Val Asp Val Arg Gly Asn Asp Phe Glu Val		
275	280	285
Ile Pro Phe Gly Ala Gly Arg Arg Ile Cys Ala Gly Met Ser Leu Gly		
290	295	300
Leu Arg Met Val Gln Leu Ser Thr Ala Thr Leu Val His Ser Phe Asn		
305	310	315 320
Trp Asp Leu Leu Asn Gly Met Ser Pro Asp Lys Leu Asp Met Glu Glu		
325	330	335
Ala Tyr Gly Leu Thr Leu Gln Arg Ala Ser Pro Leu Ile Val His Pro		
340	345	350
Lys Pro Arg Leu Ala Ser Ser Met Tyr Val Lys		
355	360	

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1757 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 35..1522

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

- 190 -

CCGTTGCTGT CGAGAAAACA GAAAGAAGAG AAAA ATG GAC TAC GTG AAT ATT 52
Met Asp Tyr Val Asn Ile
1 5

TTG CTG GGA CTG TTT TTC ACT TGG TTC TTG GTG AAT GGA CTC ATG TCA 100
Leu Leu Gly Leu Phe Phe Thr Trp Phe Leu Val Asn Gly Leu Met Ser
10 15 20

CTT CGA AGA AGA AAA ATC TCT AAG AAA CTT CCA CCA GGT CCA TTT CCT 148
Leu Arg Arg Arg Lys Ile Ser Lys Lys Leu Pro Pro Gly Pro Phe Pro
25 30 35

TTG CCT ATC ATC GGA AAT CTT CAC TTA CTT GGT AAT CAT CCT CAC AAA 196
Leu Pro Ile Ile Gly Asn Leu His Leu Leu Gly Asn His Pro His Lys
40 45 50

TCA CTT GCT CAA CTT GCA AAA ATT CAT GGT CCT ATT ATG AAT CTC AAA 244
Ser Leu Ala Gln Leu Ala Lys Ile His Gly Pro Ile Met Asn Leu Lys
55 60 65 70

TTA GGC CAA CTA AAC ACA GTG GTC ATT TCA TCA TCA GTC GTG GCA AGA 292
Leu Gly Gln Leu Asn Thr Val Val Ile Ser Ser Ser Val Val Ala Arg
75 80 85

GAA GTC TTG CAA AAA CAA GAC TTA ACA TTT TCC AAT AGG TTT GTC CCG 340
Glu Val Leu Gln Lys Gln Asp Leu Thr Phe Ser Asn Arg Phe Val Pro
90 95 100

GAC GTA GTC CAT GTC CGA AAT CAC TCC GAT TTT TCT GTT GTT TGG TTA 388
Asp Val Val His Val Arg Asn His Ser Asp Phe Ser Val Val Trp Leu
105 110 115

CCA GTC AAT TCT CGA TGG AAA ACG CTT CGC AAA ATC ATG AAC TCT AGC 436
Pro Val Asn Ser Arg Trp Lys Thr Leu Arg Lys Ile Met Asn Ser Ser
120 125 130

ATC TTT TCT GGT AAC AAG CTT GAT GGT AAT CAA CAT CTG AGG TCT AAA 484

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

- 192 -

280	285	290	
CTG GAC TTG TTT GCA GCA GGG ACT GAT ACT ACA TCG AAT ACC TTG GAG			964
Leu Asp Leu Phe Ala Ala Gly Thr Asp Thr Thr Ser Asn Thr Leu Glu			
295	300	305	310
TGG GCA ATG GCA GAA CTA CTT CAG AAT CCA CAC ACA TTG CAG AAA GCA			1012
Trp Ala Met Ala Glu Leu Leu Gln Asn Pro His Thr Leu Gln Lys Ala			
315	320		325
CAA GAA GAA CTT GCA CAA GTC ATT GGT AAA GGC AAA CAA GTA GAA GAA			1060
Gln Glu Glu Leu Ala Gln Val Ile Gly Lys Gly Lys Gln Val Glu Glu			
330	335		340
GCA GAT GTT GGA CGA CTA CCT TAC TTG CGA TGC ATA GTG AAA GAA ACC			1108
Ala Asp Val Gly Arg Leu Pro Tyr Leu Arg Cys Ile Val Lys Glu Thr			
345	350		355
TTA CGA ATA CAC CCA GCG GCT CCT CTC TTA ATT CCA CGT AAA GTG GAG			1156
Leu Arg Ile His Pro Ala Ala Pro Leu Leu Ile Pro Arg Lys Val Glu			
360	365		370
GAA GAC GTT GAG TTG TCT ACC TAT ATT ATT CCA AAG GAT TCA CAA GTT			1204
Glu Asp Val Glu Leu Ser Thr Tyr Ile Ile Pro Lys Asp Ser Gln Val			
375	380	385	390
CTA GTG AAC GTA TGG GCA ATT GGA CGC AAC TCT GAT CTA TGG GAA AAT			1252
Leu Val Asn Val Trp Ala Ile Gly Arg Asn Ser Asp Leu Trp Glu Asn			
395	400		405
CCT TTG GTC TTT AAG CCA GAA AGG TTT TGG GAG TCA GAA ATA GAT ATC			1300
Pro Leu Val Phe Lys Pro Glu Arg Phe Trp Glu Ser Glu Ile Asp Ile			
410	415		420
CGA GGT CGA GAT TTT GAA CTC ATT CCA TTT GGT GCT GGT CGA AGA ATT			1348
Arg Gly Arg Asp Phe Glu Leu Ile Pro Phe Gly Ala Gly Arg Arg Ile			
425	430	435	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

- 193 -

TGC CCT GGA TTG CCT TTG GCT ATG AGG ATG ATT CCA GTA GCA CTA GGT	1396
Cys Pro Gly Leu Pro Leu Ala Met Arg Met Ile Pro Val Ala Leu Gly	
440 445 450	
 TCA TTG CTA AAC TCA TTT AAT TGG AAA CTA TAT GGT GGA ATT GCA CCT	1444
Ser Leu Leu Asn Ser Phe Asn Trp Lys Leu Tyr Gly Gly Ile Ala Pro	
455 460 465 470	
 AAA GAT TTG GAC ATG CAG GAA AAG TTT GGC ATT ACC TTG GCG AAA GCC	1492
Lys Asp Leu Asp Met Gln Glu Lys Phe Gly Ile Thr Leu Ala Lys Ala	
475 480 485	
 CAA CCT CTG CTA GCT ATC CCA ACT CCC CTG TAGCTATAGG GATAAATTAA	1542
Gln Pro Leu Leu Ala Ile Pro Thr Pro Leu	
490 495	
 GTTGAGGTTT TAAGTTACTA GTAGATTCTA TTGCAGCTAT AGGATTTCTT TCACCATCAC	1602
GTATGCTTTA CCGTTGGATG ATGGAAAGAA ATATCTATAG CTTTGGGTTT GTTTAGTTTG	1662
CACATAAAAA TTGAATGAAT GGAATACCAT GGAGTTATAA GAAATAATAA GACTATGATT	1722
CTTACCCTAC TTGAACAATG ACATGGCTAT TTCAC	1757

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

- 194 -

TTTTTTTTTT TTTTTTA

18

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTTTTTTTTT TTTTTTC

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTTTTTTTTT TTTTTTG

18

- 195 -

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Trp Ala Ile Gly Arg Asp Pro

5

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TGGGCIATIG GI(A/C)GIGA(T/C)CC

- 196 -

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Phe Arg Pro Glu Arg Phe

5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 acids

(B) TYPE: nucleic acids

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGGAATT(T/C) (A/C) G ICCIGA(A/G) (A/C) GI TT

22

- 197 -

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CCITT(T/C)GGIG CIGGI(A/C)GI(A/C)G IATITG(T/G) (C/G) CI GG

32

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Glu Phe Xaa Pro Glu Arg Phe

- 198 -

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GAITT(T/C)IIIC CIGAI(A/C)GITT

20

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CCACACGAGT AGTTTTGGCA TTTGACCC

28

- 199 -

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GTCTTGACA TCACACTTCA ATCTG

25

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCGAATTCCC CCCCCC

17

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 nucleic acids

- 200 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCIGG(A/G)CAIA TIC(G/T) (C/T) (C/T)TICC IGCICC(A/G)AAI GG

32

SEQUENCE INFORMATION